MATHEMATISCHE ANNALEN

BEGRÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN

UNTER MITWIRKUNG

VON

LUDWIG BIEBERBACH, HARALD BOHR, MAX BORN, L. E. J. BROUWER, RICHARD COURANT, CONSTANTIN CARATHÉODORY, WALTHER V. DYCK, OTTO HÖLDER, THEODOR V. KÁRMÁN, CARL NEUMANN, ARNOLD SOMMERFELD

HERAUSGEGEBEN

VON

FELIX KLEIN
IN GÖTTINGEN

DAVID HILBERT
IN GÖTTINGEN

ALBERT EINSTEIN
IN BERLIN

OTTO BLUMENTHAL
IN AACHEN.

92. BAND

BERLIN
VERLAG VON JULIUS SPRINGER
1924

MATHEMATISCHE ANNALEN

BEGRÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN

UNTER MITWIRKUNG

VON

HERAUSGEGEBEN

VON

FELIX KLEIN
IN GÖTTINGEN

DAVID HILBERT
IN GÖTTINGEN

ALBERT EINSTEIN
IN BERLIN

OTTO BLUMENTHAL
IN AACHEN.

92. BAND

BERLIN
VERLAG VON JULIUS SPRINGER
1924

Inhalt des zweiundneunzigsten Bandes.

(In alphabetischer Ordnung.)

Die Grundlagen der Physik.

Von

David Hilbert in Göttingen.

Das Nachfolgende ist im wesentlichen ein Abdruck der beiden älteren Mitteilungen[1]) von mir über die „Grundlagen der Physik" und meiner Bemerkungen dazu, die F. Klein in seiner Mitteilung[2]) „Zu Hilberts erster Note über die Grundlagen der Physik" veröffentlicht hat — mit nur geringfügigen redaktionellen Abweichungen und Umstellungen, die das Verständnis erleichtern sollen.

Das mechanistische Einheitsideal in der Physik, wie es von den großen Forschern der vorangegangenen Generation geschaffen und noch während der Herrschaft der klassischen Elektrodynamik festgehalten worden war, muß heute endgültig aufgegeben werden. Durch die Aufstellung und Entwickelung des Feldbegriffes bildete sich allmählich eine neue Möglichkeit für die Auffassung der physikalischen Welt aus. Mie zeigte als der erste einen Weg, auf dem dieses neuentstandene „feldtheoretische Einheitsideal", wie ich es nennen möchte, der allgemeinen mathematischen Behandlung zugänglich gemacht werden kann. Während die alte mechanistische Auffassung unmittelbar die Materie selbst als Ausgang nimmt und diese durch eine endliche Auswahl diskreter Parameter bestimmt ansetzt, dient vielmehr dem neuen feldtheoretischen Ideal das physikalische Kontinuum, die sogenannte Raum-Zeit-Mannigfaltigkeit, als Fundament. Waren früher Differentialgleichungen mit einer unabhängigen Variablen die Form der Weltgesetze, so sind jetzt notwendig partielle Differentialgleichungen ihre Ausdrucksform.

Die gewaltigen Problemstellungen und Gedankenbildungen der allgemeinen Relativitätstheorie von Einstein finden nun, wie ich in meiner ersten Mitteilung ausgeführt habe, auf dem von Mie betretenen Wege

---

[1]) Göttinger Nachr.: Erste Mitteilung, vorgelegt am 20. Nov. 1915, zweite Mitteilung, vorgelegt am 23. Dez. 1916.

[2]) Göttinger Nachr.: vorgelegt am 25. Jan. 1918.

ihren einfachsten und natürlichsten Ausdruck und zugleich in formaler Hinsicht eine systematische Ergänzung und Abrundung.

Seit der Veröffentlichung meiner ersten Mitteilung sind bedeutsame Abhandlungen über diesen Gegenstand erschienen: ich erwähne nur die glänzenden und tiefsinnigen Untersuchungen von Weyl und die an immer neuen Ansätzen und Gedanken reichen Mitteilungen von Einstein. Indes sowohl Weyl gibt späterhin seinem Entwicklungsgange eine solche Wendung, daß er auf die von mir aufgestellten Gleichungen ebenfalls gelangt, und andererseits auch Einstein, obwohl wiederholt von abweichenden und unter sich verschiedenen Ansätzen ausgehend, kehrt schließlich in seinen letzten Publikationen geradenwegs zu den Gleichungen meiner Theorie zurück.

Ich glaube sicher, daß die hier von mir entwickelte Theorie einen bleibenden Kern enthält und einen Rahmen schafft, innerhalb dessen für den künftigen Aufbau der Physik im Sinne eines feldtheoretischen Einheitsideals genügender Spielraum da ist. Auch ist es auf jeden Fall von erkenntnistheoretischem Interesse, zu sehen, wie die wenigen einfachen in den Axiomen I, II, III, IV von mir ausgesprochenen Annahmen zum Aufbau der ganzen Theorie genügend sind.

Ob freilich das reine feldtheoretische Einheitsideal ein definitives ist, evtl. welche Ergänzungen und Modifikationen desselben nötig sind, um insbesondere die theoretische Begründung für die Existenz des negativen und des positiven Elektrons, sowie den widerspruchsfreien Aufbau der im Atominneren geltenden Gesetze zu ermöglichen, — dies zu beantworten, ist die Aufgabe der Zukunft.

Teil I.

Es seien x_s $(s=1, 2, 3, 4)$ irgendwelche die Weltpunkte wesentlich eindeutig benennende Koordinaten, die sogenannten Weltparameter (allgemeinste Raum-Zeit-Koordinaten). Die das Geschehen in x_s charakterisierenden Größen seien:

1. die zuerst von Einstein eingeführten Gravitationspotentiale $g_{\mu\nu}$ $(\mu, \nu = 1, 2, 3, 4)$ mit symmetrischem Tensorcharakter gegenüber einer beliebigen Transformation der Weltparameter x_s; sie bilden die Koeffizienten der invarianten Differentialform

$$\sum_{\mu\nu} g_{\mu\nu}\, dx_\mu\, dx_\nu\,;$$

2. die vier elektrodynamischen Potentiale q_s mit Vektorcharakter im selben Sinne, welche die Koeffizienten der invarianten Linearform

$$\sum_s q_s\, dx_s$$

bilden.

Das physikalische Geschehen ist nicht willkürlich, es gelten vielmehr folgende Axiome:

Axiom I (Mies Axiom von der Weltfunktion[3]). *Das Gesetz des physikalischen Geschehens bestimmt sich durch eine Weltfunktion H, die folgende Argumente enthält:*

$$(1)\qquad g_{\mu\nu},\; g_{\mu\nu l}=\frac{\partial g_{\mu\nu}}{\partial x_l},\qquad g_{\mu\nu lk}=\frac{\partial^2 g_{\mu\nu}}{\partial x_l\,\partial x_k},$$

$$(2)\qquad q_s,\; q_{sl}=\frac{\partial q_s}{\partial x_l},\qquad (s,l=1,2,3,4)$$

und zwar muß die Variation des Integrals

$$\int H\sqrt{g}\,d\omega$$

$$(g=-|g_{\mu\nu}|,\qquad d\omega=dx_1\,dx_2\,dx_3\,dx_4)$$

für jedes der 14 Potentiale $g_{\mu\nu}$, q_s verschwinden.

An Stelle der Argumente (1) können offenbar auch die Argumente

$$(3)\qquad g^{\mu\nu},\; g^{\mu\nu}_l=\frac{\partial g^{\mu\nu}}{\partial x_l},\qquad g^{\mu\nu}_{lk}=\frac{\partial^2 g^{\mu\nu}}{\partial x_l\,\partial x_k}$$

treten, wobei $g^{\mu\nu}$ die durch $(-g)$ dividierte Unterdeterminante der Determinante $(-g)$ in bezug auf ihr Element $g_{\mu\nu}$ bedeutet.

Aus Axiom I folgen zunächst bezüglich der zehn Gravitationspotentiale $g^{\mu\nu}$ die zehn Lagrangeschen Differentialgleichungen

$$(4)\qquad \frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}}-\sum_k\frac{\partial}{\partial x_k}\frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}_k}+\sum_{k,l}\frac{\partial^2}{\partial x_k\,\partial x_l}\frac{\partial\sqrt{g}H}{\partial g^{\mu\nu}_{kl}}=0\quad(\mu,\nu=1,2,3,4),$$

und sodann bezüglich der vier elektrodynamischen Potentiale q_s die vier Lagrangeschen Differentialgleichungen

$$(5)\qquad \frac{\partial\sqrt{g}H}{\partial q_h}-\sum_k\frac{\partial}{\partial x_k}\frac{\partial\sqrt{g}H}{\partial q_{hk}}=0,\qquad (h=1,2,3,4).$$

Bezüglich der Differentialquotienten nach $g^{\mu\nu}$, $g^{\mu\nu}_k$, $g^{\mu\nu}_{kl}$, wie sie in (4) und nachfolgenden Formeln auftreten, sei ein für allemal bemerkt, daß wegen der Symmetrie in μ,ν einerseits und k,l andererseits die Differentialquotienten nach $g^{\mu\nu}$, $g^{\mu\nu}_k$ so zu verstehen sind, daß man ihnen den Faktor 1 bzw. $\frac{1}{2}$ hinzusetzt, je nachdem $\mu=\nu$ bzw. $\mu\neq\nu$ ausfällt, ferner die Differentialquotienten nach $g^{\mu\nu}_{kl}$ mit 1 bzw. $\frac{1}{2}$ bzw. $\frac{1}{4}$ multipliziert zu

[3]) Mies Weltfunktionen enthalten nicht genau diese Argumente; insbesondere geht der Gebrauch der Argumente (2) auf Born zurück; es ist jedoch gerade die Einführung und Verwendung einer solchen Weltfunktion im Hamiltonschen Prinzip das Charakteristische der Mieschen Elektrodynamik.

nehmen sind, je nachdem $\mu = \nu$ und $k = l$ bzw. $\mu = \nu$ und $k \neq l$ oder $\mu \neq \nu$ und $k = l$ bzw. $\mu \neq \nu$ und $k \neq l$ ausfällt.

Der Kürze halber bezeichnen wir die linken Seiten der Gleichungen (4), (5) bzw. mit

$$[\sqrt{g}H]_{\mu\nu}, \qquad [\sqrt{g}H]_{h}.$$

Die Gleichungen (4) mögen die Grundgleichungen der Gravitation, die Gleichungen (5) die elektrodynamischen Grundgleichungen heißen.

Axiom II (Axiom von der allgemeinen Invarianz[4]). *Die Weltfunktion H ist eine Invariante gegenüber einer beliebigen Transformation der Weltparameter x_s.*

Axiom II ist der einfachste mathematische Ausdruck für die Forderung, daß die Koordinaten an sich keinerlei physikalische Bedeutung haben, sondern nur eine Numerierung der Weltpunkte darstellen, von deren Art die Verkettung der Potentiale $g_{\mu\nu}$, q_s völlig unabhängig ist.

Im folgenden benutzen wir die leicht beweisbare Tatsache, daß, wenn p^j $(j = 1, 2, 3, 4)$ einen willkürlichen kontravarianten Vektor bedeutet, der Ausdruck

$$p^{\mu\nu} = \sum_s (g_s^{\mu\nu} p^s - g^{\mu s} p_s^\nu - g^{\nu s} p_s^\mu), \qquad \left(p_s^j = \frac{\partial p^j}{\partial x_s}\right)$$

einen symmetrischen kontravarianten Tensor und der Ausdruck

$$p_l = \sum_s (q_{ls} p^s + q_s p_l^s) \text{ [5]}$$

einen kovarianten Vektor darstellt.

Des weiteren stellen wir zwei mathematische Theoreme auf, die wie folgt lauten:

Theorem 1. Wenn J eine von $g^{\mu\nu}$, $g_l^{\mu\nu}$, $g_{lk}^{\mu\nu}$, q_s, q_{sk} abhängige Invariante ist, so gilt stets identisch in allen Argumenten und für jeden willkürlichen kontravarianten Vektor p^s

$$\sum_{\mu,\nu,l,k} \left(\frac{\partial J}{\partial g^{\mu\nu}} \Delta g^{\mu\nu} + \frac{\partial J}{\partial g_l^{\mu\nu}} \Delta g_l^{\mu\nu} + \frac{\partial J}{\partial g_{lk}^{\mu\nu}} \Delta g_{lk}^{\mu\nu}\right) + \sum_{s,k} \left(\frac{\partial J}{\partial q_s} \Delta q_s + \frac{\partial J}{\partial q_{sk}} \Delta q_{sk}\right) = 0;$$

[4]) Die Forderung der orthogonalen Invarianz hat bereits Mie gestellt. In dem oben aufgestellten Axiom II findet der Einsteinsche fundamentale Grundgedanke der allgemeinen Invarianz den einfachsten Ausdruck, wennschon bei Einstein das Hamiltonsche Prinzip nur eine Nebenrolle spielt und seine Funktionen H keineswegs allgemeine Invarianten sind, auch die elektrischen Potentiale nicht enthalten.

[5]) p_l ist nicht zu verwechseln mit dem zu p^s gehörigen kovarianten Vektor $\sum_s g_{ls} p^s$.

dabei ist

$$\begin{aligned}
\Delta g^{\mu\nu} &= \sum_m (g^{\mu m} p_m^\nu + g^{\nu m} p_m^\mu),\\
\Delta g_l^{\mu\nu} &= -\sum_m g_m^{\mu\nu} p_l^m + \frac{\partial \Delta g^{\mu\nu}}{\partial x_l},\\
\Delta g_{lk}^{\mu\nu} &= -\sum_m (g_m^{\mu\nu} p_{lk}^m + g_{lm}^{\mu\nu} p_k^m + g_{km}^{\mu\nu} p_l^m) + \frac{\partial^2 \Delta g^{\mu\nu}}{\partial x_l \partial x_k},\\
\Delta q_s &= -\sum_m q_m p_s^m,\\
\Delta q_{sk} &= -\sum_m q_{sm} p_k^m + \frac{\partial \Delta q_s}{\partial x_k}.
\end{aligned}$$

Dieses Theorem 1 läßt sich auch folgendermaßen aussprechen:

Wenn J eine Invariante und p^s ein willkürlicher Vektor wie vorhin ist, so gilt die Identität

$$\sum_s \frac{\partial J}{\partial x_s} p^s = P(J); \tag{6}$$

dabei ist

$$P = P_g + P_q,$$

$$P_g = \sum_{\mu,\nu,l,k} \left(p^{\mu\nu} \frac{\partial}{\partial g^{\mu\nu}} + p_l^{\mu\nu} \frac{\partial}{\partial g_l^{\mu\nu}} + p_{lk}^{\mu\nu} \frac{\partial}{\partial g_{lk}^{\mu\nu}} \right),$$

$$P_q = \sum_{l,k} \left(p_l \frac{\partial}{\partial q_l} + p_{lk} \frac{\partial}{\partial q_{lk}} \right)$$

gesetzt und es gelten die Abkürzungen:

$$p_k^{\mu\nu} = \frac{\partial p^{\mu\nu}}{\partial x_k}, \qquad p_{kl}^{\mu\nu} = \frac{\partial^2 p^{\mu\nu}}{\partial x_k \partial x_l}, \qquad p_{lk} = \frac{\partial p_l}{\partial x_k}.$$

Der Beweis von (6) ergibt sich leicht; denn diese Identität ist offenbar richtig, wenn p^s ein konstanter Vektor ist, und daraus folgt sie wegen ihrer Invarianz allgemein.

Theorem 2. Wenn J, wie im Theorem 1, eine von $g^{\mu\nu}$, $g_l^{\mu\nu}$, $g_{lk}^{\mu\nu}$, q_s, q_{sk} abhängige Invariante ist, und, wie oben, die Variationsableitungen von $\sqrt{g}J$ bez. $g^{\mu\nu}$ mit $[\sqrt{g}J]_{\mu\nu}$, bez. q_μ mit $[\sqrt{g}J]_\mu$ bezeichnet werden, und wenn ferner zur Abkürzung:

$$i_s = \sum_{\mu,\nu} ([\sqrt{g}J]_{\mu\nu} g_s^{\mu\nu} + [\sqrt{g}J]_\mu q_{\mu s}),$$

$$i_s^l = -2 \sum_\mu [\sqrt{g}J]_{\mu s} g^{\mu l} + [\sqrt{g}J]_l q_s,$$

gesetzt wird, so gelten die Identitäten

$$i_s = \sum_l \frac{\partial i_s^l}{\partial x_l} \qquad (s = 1, 2, 3, 4). \tag{7}$$

Dieses Theorem 2 enthält als wesentlichen Kern einen allgemeinen mathematischen Satz*), der mir das Leitmotiv für den Aufbau der Theorie gewesen ist und der sich folgendermaßen ausspricht:

Ist F eine Funktion von n Größen (Funktionen von x_1, x_2, x_3, x_4) und ihren Ableitungen, und ist das Integral

$$\int F d\omega$$

invariant bei beliebigen Transformationen der vier Weltparameter x_1, x_2, x_3, x_4, so sind in dem System der n Lagrangeschen Differentialgleichungen, welche zu dem Variationsproblem

$$\delta \int F d\omega = 0$$

gehören, stets vier eine Folge der $n-4$ übrigen in dem Sinne, daß zwischen den n Lagrangeschen Ableitungen von F in bezug auf jene n Größen und deren totalen Differentialquotienten nach x_1, x_2, x_3, x_4 stets vier linear unabhängige Relationen identisch erfüllt sind.

Zum Beweise von Theorem 2 betrachten wir ein endliches Stück der vierdimensionalen Welt; ferner sei p^s ein Vektor, der nebst seinen Ableitungen auf der dreidimensionalen Oberfläche jenes Weltstückes verschwindet. Man hat gemäß der Definition von P:

$$P(\sqrt{g}J) = \sqrt{g}P(J) + J\sum_{\mu,\nu}\frac{\partial\sqrt{g}}{\partial g^{\mu\nu}}p^{\mu\nu} = \sqrt{g}P(J) + J\sum_{s}\left(\frac{\partial\sqrt{g}}{\partial x_s}p^s + \sqrt{g}p^s_s\right),$$

und nach Theorem 1 daher:

$$P(\sqrt{g}J) = \sqrt{g}\sum_{s}\frac{\partial J}{\partial x_s}p^s + J\sum_{s}\left(\frac{\partial\sqrt{g}}{\partial x_s}p^s + \sqrt{g}p^s_s\right) = \sum_{s}\frac{\partial\sqrt{g}Jp^s}{\partial x_s}.$$

Integrieren wir diese Gleichung über das betrachtete Weltstück, so ergibt sich wegen der Divergenzform der rechten Seite und wegen der Annahme über p^s:

$$\int P(\sqrt{g}J)\,d\omega = 0.$$

Wegen der Bildungsweise der Lagrangeschen Ableitung ist demnach auch

$$\int\left\{\sum_{\mu,\nu}[\sqrt{g}J]_{\mu\nu}p^{\mu\nu} + \sum_{\mu}[\sqrt{g}J]_{\mu}p_{\mu}\right\}d\omega = 0.$$

*) Den Beweis dieses Satzes hat allgemein Emmy Noether geliefert (Göttinger Nachr. 1918, Heft 2: „Invariante Variationsprobleme"). Die in Theorem 2 angegebenen Identitäten sind in meiner ersten Mitteilung zwar nur für den Fall behauptet worden, daß die Invariante von den $g^{\mu\nu}$ und deren Ableitungen abhängt; aber das dort eingeschlagene und im Text reproduzierte Beweisverfahren gilt ebenso auch für unsere allgemeine Invariante J. In der allgemeinen Form sind die angegebenen Identitäten zuerst von F. Klein auf Grund der Methode der infinitesimalen Transformation abgeleitet worden (Gött. Nachr. 1917, Heft 3: „Zu Hilberts erster Note über die Grundlagen der Physik").

Der Integrand hier läßt sich in der Form

$$\sum_{s,l} (i_s p^s + i_s^l p_l^s)$$

schreiben. Aus der so entstehenden Formel

$$\int \sum_{s,l} (i_s p^s + i_s^l p_l^s)\, d\omega = 0$$

erhalten wir

$$\sum_s \int \left(i_s - \sum_l \frac{\partial i_s^l}{\partial x_l} \right) p^s\, d\omega = 0$$

und damit auch die Behauptung unseres Theorems 2.

Zur Bestimmung der Weltfunktion H sind noch weitere Axiome erforderlich. Sollen die Grundgleichungen (4), (5) der Gravitation und der Elektrodynamik nur zweite Ableitungen der $g^{\mu\nu}$ enthalten, so muß H sich additiv zusammensetzen aus einer linearen Funktion mit konstanten Koeffizienten von der Invariante

$$K = \sum_{\mu,\nu} g^{\mu\nu} K_{\mu\nu},$$

wo $K_{\mu\nu}$ den Riemannschen Krümmungstensor

$$K_{\mu\nu} = \sum_{\varkappa} \left(\frac{\partial}{\partial x_\nu} \left\{ {\mu\varkappa \atop \varkappa} \right\} - \frac{\partial}{\partial x_\varkappa} \left\{ {\mu\nu \atop \varkappa} \right\} \right) + \sum_{\varkappa,\lambda} \left(\left\{ {\mu\varkappa \atop \lambda} \right\} \left\{ {\lambda\nu \atop \varkappa} \right\} - \left\{ {\mu\nu \atop \lambda} \right\} \left\{ {\lambda\varkappa \atop \varkappa} \right\} \right)$$

bedeutet, und einer Invariante L, die nur von $g^{\mu\nu}$, $g_l^{\mu\nu}$, q_s, q_{sk} abhängt. Wir machen folgende spezielle Annahme:

Axiom III (Axiom von der Gravitation und der Elektrizät). *Die Weltfunktion H hat die Gestalt*

$$H = K + L,$$

wo K die aus dem Riemannschen Tensor entspringende Invariante, die Krümmung, ist und L nur von den $g^{\mu\nu}$, q_s, q_{sk} abhängt.

Hiernach nehmen die Gravitationsgleichungen die Form

$$(8) \qquad [\sqrt{g}K]_{\mu\nu} = -\frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} \qquad (\mu, \nu = 1, 2, 3, 4),$$

und die elektrodynamischen Gleichungen die Form

$$(9) \qquad \frac{1}{\sqrt{g}} \sum_k \frac{\partial}{\partial x_k} \frac{\partial \sqrt{g} L}{\partial q_{hk}} = \frac{\partial L}{\partial q_h} \qquad (h = 1, 2, 3, 4)$$

an.

Um den Ausdruck von $[\sqrt{g}K]_{\mu\nu}$ zu bestimmen, spezialisiere man zunächst das Koordinatensystem so, daß für den betrachteten Weltpunkt die $g_s^{\mu\nu}$ sämtlich verschwinden. Man findet auf diese Weise:

$$[\sqrt{g}K]_{\mu\nu} = \sqrt{g}\left(K_{\mu\nu} - \frac{1}{2} g_{\mu\nu} K\right).$$

Führen wir noch für den Tensor

$$-\frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}$$

die Bezeichnung $T_{\mu\nu}$ ein, so lauten die Gravitationsgleichungen

$$K_{\mu\nu}-\frac{1}{2}g_{\mu\nu}K=T_{\mu\nu}.$$

Andererseits wenden wir auf die Invariante L das Theorem 1 an und erhalten dadurch

$$(10)\qquad \sum_{\mu,\nu,m}\frac{\partial L}{\partial g^{\mu\nu}}(g^{\mu m}p^{\nu}_{m}+g^{\nu m}p^{\mu}_{m})-\sum_{s,m}\frac{\partial L}{\partial q_s}q_m p^m_s$$
$$-\sum_{s,k,m}\frac{\partial L}{\partial q_{sk}}(q_{sm}p^m_k+q_{mk}p^m_s+q_m p^m_{sk})=0.$$

Das Nullsetzen des Koeffizienten von p^m_{sk} linker Hand liefert die Gleichung

$$\left(\frac{\partial L}{\partial q_{sk}}+\frac{\partial L}{\partial q_{ks}}\right)q_m=0$$

oder

$$(11)\qquad \frac{\partial L}{\partial q_{sk}}+\frac{\partial L}{\partial q_{ks}}=0,$$

d. h. die Ableitungen der elektrodynamischen Potentiale q_s treten nur in den Verbindungen

$$M_{ks}=q_{sk}-q_{ks}$$

auf. Damit erkennen wir, daß bei unseren Annahmen die Invariante L außer von den Potentialen $g^{\mu\nu}$, q_s lediglich von den Komponenten des schiefsymmetrischen invarianten Tensors

$$M=(M_{ks})=\mathrm{Rot}(q_s),$$

d. h. des sogenannten elektromagnetischen Sechservektors abhängt. Und hieraus folgt weiter, daß

$$\frac{\partial L}{\partial q_{sk}}=\frac{\partial L}{\partial M_{ks}}=H^{ks}$$

ein schiefsymmetrischer kontravarianter Tensor, sowie daß

$$\frac{\partial L}{\partial q_k}=r^k$$

ein kontravarianter Vektor ist.

Mit Anwendung der eingeführten Bezeichnungen erhalten die elektrodynamischen Gleichungen die Form:

$$(12)\qquad \frac{1}{\sqrt{g}}\sum_k\frac{\partial\sqrt{g}H^{kh}}{\partial x_k}=r^h \qquad (h=1,2,3,4).$$

Man erkennt in diesen Gleichungen eine Verallgemeinerung des einen Systems der Maxwellschen Gleichungen; das andere System erhält man aus den Gleichungen:

$$M_{ks} = q_{sk} - q_{ks}$$

durch Differentiation und Addition:

$$\frac{\partial M_{ks}}{\partial x_t} + \frac{\partial M_{st}}{\partial x_k} + \frac{\partial M_{tk}}{\partial x_s} = 0 \qquad (t, k, s = 1, 2, 3, 4). \tag{13}$$

Wir sehen also, daß die Form dieser „verallgemeinerten Maxwellschen Gleichungen“ (12), (13) im wesentlichen schon durch die Forderung der allgemeinen Invarianz, also auf Grund von Axiom II, bestimmt ist. Setzen wir in der Identität (10) den Koeffizienten von p_m^ν linker Hand gleich Null, so erhalten wir mit Benutzung von (11)

$$2\sum_{\mu}\frac{\partial L}{\partial g^{\mu\nu}}g^{\mu m} - \frac{\partial L}{\partial q_m}q_\nu - \sum_{s}\frac{\partial L}{\partial M_{ms}}M_{\nu s} = 0 \qquad (\mu = 1, 2, 3, 4), \tag{14}$$

also

$$2\sum_{\mu}\frac{\partial L}{\partial g^{\mu\nu}}g^{\mu m} = \sum_{s}H^{ms}M_{\nu s} + r^m q_\nu$$

oder

$$-\frac{2}{\sqrt{g}}\sum_{\mu}\frac{\partial\sqrt{g}L}{\partial g^{\mu\nu}}g^{\mu m} = L\delta_\nu^m - \sum_{s}H^{ms}M_{\nu s} - r^m q_\nu,$$

$$\delta_\nu^m = 0 \quad (m \neq \nu),$$

$$\delta_\nu^\nu = 1.$$

Demnach ergibt sich für $T_{\mu\nu}$ die Darstellung:

$$T_{\mu\nu} = \sum_{\mu} g_{\mu m}T_\nu^m,$$

$$T_\nu^m = \frac{1}{2}\{L\delta_\nu^m - \sum_{s}H^{ms}M_{\nu s} - r^m q_\nu\}.$$

Der Ausdruck rechts stimmt überein mit dem Mieschen elektromagnetischen Energietensor, und wir finden also, daß der Miesche Energietensor nichts anderes ist als der durch Differentiation der Invariante L nach den Gravitationspotentialen $g^{\mu\nu}$ entstehende allgemein invariante Tensor — ein Umstand, der mich zum erstenmal auf den notwendigen engen Zusammenhang zwischen der Einsteinschen allgemeinen Relativitätstheorie und der Mieschen Elektrodynamik hingewiesen und mir die Überzeugung von der Richtigkeit der hier entwickelten Theorie gegeben hat.

Die Anwendung des Theorems 2 auf die Invariante K liefert:

$$\sum_{\mu,\nu}[\sqrt{g}K]_{\mu\nu}g_s^{\mu\nu} + 2\sum_{m}\frac{\partial}{\partial x_m}\Big(\sum_{\mu}[\sqrt{g}K]_{\mu s}g^{\mu m}\Big) = 0. \tag{15a}$$

Die Anwendung auf L ergibt

(15b)
$$\sum_{\mu,\nu}(-\sqrt{g}\,T_{\mu\nu})g_s^{\mu\nu}+2\sum_m\frac{\partial}{\partial x_m}(-\sqrt{g}\,T_s^m)$$
$$+\sum_\mu[\sqrt{g}\,L]_\mu q_{\mu s}-\sum_\mu\frac{\partial}{\partial x_\mu}([\sqrt{g}\,L]_\mu q_s)=0\qquad (s=1,2,3,4).$$

Als Folge der elektrodynamischen Grundgleichungen erhalten wir hieraus:

(16)
$$\sum_{\mu,\nu}\sqrt{g}\,T_{\mu\nu}g_s^{\mu\nu}+2\sum_m\frac{\partial\sqrt{g}\,T_s^m}{\partial x_m}=0.$$

Diese Gleichungen (16) ergeben sich auch als Folge der Gravitationsgleichungen, auf Grund von (15a). Sie haben die Bedeutung der mechanischen Grundgleichungen. Im Falle der speziellen Relativität, wenn die $g_{\mu\nu}$ Konstante sind, gehen sie über in die Gleichungen

$$\sum\frac{\partial T_s^m}{\partial x_m}=0,$$

welche die Erhaltung von Energie und Impuls ausdrücken.

Aus den Gleichungen (16) folgt auf Grund der Identitäten (15b):

$$\sum_\mu[\sqrt{g}\,L]_\mu q_{\mu s}-\sum_\mu\frac{\partial}{\partial x_\mu}([\sqrt{g}\,L]_\mu q_s)=0$$

oder

(17)
$$\sum_\mu\left\{M_{\mu s}[\sqrt{g}\,L]_\mu+q_s\frac{\partial}{\partial x_\mu}([\sqrt{g}\,L]_\mu)\right\}=0,$$

d. h. aus den Gravitationsgleichungen (4) folgen vier voneinander unabhängige lineare Relationen zwischen den elektrodynamischen Grundgleichungen (5) und ihren ersten Ableitungen. Dies ist der genaue mathematische Ausdruck für den Zusammenhang zwischen Gravitation und Elektrodynamik, der die ganze Theorie beherrscht.

Da L unserer Annahme zufolge nicht von den Ableitungen der $g^{\mu\nu}$ abhängen soll, so muß L eine Funktion von gewissen vier allgemeinen Invarianten sein, die den von Mie angegebenen speziellen orthogonalen Invarianten entsprechen und von denen die beiden einfachsten diese sind:

$$Q=\sum_{k,l,m,n}M_{mn}M_{kl}g^{mk}g^{nl}$$

und

$$q=\sum_{k,l}q_kq_lg^{kl}.$$

Der einfachste und im Hinblick auf den Bau von K nächstliegende Ansatz für L ist zugleich derjenige, der der Mieschen Elektrodynamik entspricht, nämlich

$$L=\alpha Q+f(q)\qquad(\alpha=\text{konst.}).$$

Gemäß diesem Ansatz erhält man zwischen den Größen, die in den verallgemeinerten Maxwellschen Gleichungen auftreten, die Beziehungen

$$H^{ks} = 4\alpha M^{ks},$$
$$r^k = 2f'(q)q^k,$$

wo

$$M^{ks} = \sum_{\mu,\nu} g^{k\mu} g^{s\nu} M_{\mu\nu},$$
$$q^k = \sum_l g^{kl} q_l$$

zu setzen ist. Für den ganz speziellen Fall

$$f(q) = \beta q \qquad (\beta = \text{konst.})$$

folgt, daß der „Stromvektor“ r^k proportional dem kontravarianten Vektor q^k wird.

Teil II.

Es soll nun der Zusammenhang der Theorie mit der Erfahrung näher erörtert werden. Dazu ist noch ein weiteres Axiom erforderlich.

Axiom IV (Raum-Zeit-Axiom). *Es soll die quadratische Form*

$$G(X_1, X_2, X_3, X_4) = \sum_{\mu\nu} g_{\mu\nu} X_\mu X_\nu \tag{18}$$

von der Art sein, daß bei ihrer Darstellung als Summe von vier Quadraten linearer Formen der X_s stets drei Quadrate mit positivem und ein Quadrat mit negativem Vorzeichen auftritt.

Die quadratische Form (18) liefert für unsere vierdimensionale Welt der x_s die Maßbestimmung einer Pseudogeometrie. Die Determinante g der $g_{\mu\nu}$ fällt negativ aus.

Ist in dieser Geometrie eine Kurve

$$x_s = x_s(p) \qquad (s = 1, 2, 3, 4)$$

gegeben, wo $x_s(p)$ irgendwelche reelle Funktionen des Parameters p bedeuten, so kann diese in Teilstücke zerlegt werden, auf denen einzeln der Ausdruck

$$G\left(\frac{dx_1}{dp}, \frac{dx_2}{dp}, \frac{dx_3}{dp}, \frac{dx_4}{dp}\right)$$

nicht sein Vorzeichen ändert: ein Kurvenstück, für welches

$$G\left(\frac{dx_s}{dp}\right) > 0$$

ausfällt, heiße eine *Strecke* und das längs dieses Kurvenstücks genommene Integral

$$\lambda = \int \sqrt{G\left(\frac{dx_s}{dp}\right)}\, dp$$

heiße die *Länge der Strecke*; ein Kurvenstück, für welches

$$G\left(\frac{dx_s}{dp}\right) < 0$$

ausfällt, heiße eine *Zeitlinie* und das längs dieses Kurvenstückes genommene Integral

$$\tau = \int \sqrt{-G\left(\frac{dx_s}{dp}\right)}\, dp$$

heiße die *Eigenzeit der Zeitlinie*; endlich heiße ein Kurvenstück, längs dessen

$$G\left(\frac{dx_s}{dp}\right) = 0$$

wird, eine *Nullinie*.

Um diese Begriffe unserer Pseudogeometrie anschaulich zu machen, denken wir uns ein ideales Maßinstrument: die *Lichtuhr*, mittels derer wir die Eigenzeit längs einer jeden Zeitlinie bestimmen können.

Zunächst zeigen wir, daß dieses Instrument ausreicht, um mit seiner Hilfe die Werte der $g_{\mu\nu}$ als Funktionen von x_s zu berechnen, sobald nur ein bestimmtes Raum-Zeit-Koordinatensystem x_s eingeführt worden ist. In der Tat wählen wir irgend zehn Zeitlinien aus, die sämtlich längs verschiedenen Richtungen in den nämlichen Weltpunkt x_s einlaufen, so daß diesem Endpunkt jedesmal der Parameterwert p zukommt, so ergibt sich für jede der zehn Zeitlinien im Endpunkt die Gleichung

$$\left(\frac{d\lambda^{(h)}}{dp}\right)^2 = G\left(\frac{dx_s^{(h)}}{dp}\right), \qquad (h = 1, 2, \ldots, 10);$$

hier sind die linken Seiten bekannt, sobald wir die Eigenzeiten $\tau^{(h)}$ mittels der Uhr bestimmt haben. Setzen wir nun zur Abkürzung

$$D(u) = \begin{vmatrix} \left(\frac{dx_1^{(1)}}{dp}\right)^2, & \frac{dx_1^{(1)}}{dp}\frac{dx_2^{(1)}}{dp}, & \ldots, & \left(\frac{dx_4^{(1)}}{dp}\right)^2, & \left(\frac{d\lambda^{(1)}}{dp}\right)^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \left(\frac{dx_1^{(10)}}{dp}\right)^2, & \frac{dx_1^{(10)}}{dp}\frac{dx_2^{(10)}}{dp}, & \ldots, & \left(\frac{dx_4^{(10)}}{dp}\right)^2, & \left(\frac{d\lambda^{(10)}}{dp}\right)^2 \\ X_1^2, & X_1 X_2, & \ldots, & X_4^2, & u \end{vmatrix},$$

so wird offenbar

$$G(X_s) = -\frac{D(0)}{\frac{\partial D}{\partial u}}, \tag{19}$$

wodurch sich zugleich für die Richtungen der ausgewählten zehn Zeitlinien im Punkte $x_s(p)$ die Bedingung

$$\frac{\partial D}{\partial u} \neq 0$$

als notwendig herausstellt.

Ist G nach (19) berechnet, so würde die Anwendung des Verfahrens auf irgendeine 11-te Zeitlinie, die in $x_s(p)$ endigt, die Gleichung

$$\left(\frac{d\lambda^{(11)}}{dp}\right)^2 = G\left(\frac{dx^{(11)}}{dp}\right)$$

liefern und diese Gleichung wäre dann sowohl eine Kontrolle für die Richtigkeit des Instrumentes als auch eine experimentelle Bestätigung dafür, daß die Voraussetzungen der Theorie für die wirkliche Welt zutreffen.

Der axiomatische Aufbau unserer Pseudogeometrie ließe sich ohne Schwierigkeit durchführen: erstens ist ein Axiom aufzustellen, auf Grund dessen folgt, daß Länge bzw. Eigenzeit Integrale sein müssen, deren Integrand lediglich eine Funktion der x_s und ihrer ersten Ableitungen nach dem Parameter ist; als ein solches Axiom wäre etwa der bekannte Enveloppensatz für geodätische Linien verwendbar. Zweitens ist ein Axiom erforderlich, wonach die Sätze der pseudo-Euklidischen Geometrie d. h. das alte Relativitätsprinzip im Unendlichkleinen gelten soll; hierzu wäre das von W. Blaschke[7]) aufgestellte Axiom besonders geeignet, welches aussagt, daß die Bedingung der Orthogonalität für irgend zwei Richtungen — sei es bei Strecken oder Zeitlinien — stets eine gegenseitige sein soll.

Es seien noch kurz die hauptsächlichsten Tatsachen zusammengestellt, die uns die Monge-Hamiltonsche Theorie der Differentialgleichungen für unsere Pseudogeometrie lehrt.

Jedem Weltpunkte x_s gehört ein Kegel zweiter Ordnung zu, der in x_s seine Spitze hat und in den laufenden Punktkoordinaten X_s durch die Gleichung

$$G(X_1 - x_1, X_2 - x_2, X_3 - x_3, X_4 - x_4) = 0$$

bestimmt ist; derselbe heiße der zum Punkte x_s zugehörige *Nullkegel*. Die sämtlichen Nullkegel bilden ein vierdimensionales Kegelfeld, zu dem einerseits die „Mongesche" Differentialgleichung

$$G\left(\frac{dx_1}{dp}, \frac{dx_2}{dp}, \frac{dx_3}{dp}, \frac{dx_4}{dp}\right) = 0$$

und andererseits die „Hamiltonsche" partielle Differentialgleichung

$$(20) \qquad H\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \frac{\partial f}{\partial x_4}\right) = 0$$

gehört, wo H die zu G reziproke quadratische Form

$$H(U_1, U_2, U_3, U_4) = \sum_{\mu\nu} g^{\mu\nu} U_\mu U_\nu$$

---

[7]) Räumliche Variationsprobleme mit symmetrischer Transversalitätsbedingung, Leipziger Berichte, Math.-phys. Kl. 68 (1916), S. 50.

bedeutet. Die Charakteristiken der Mongeschen und zugleich die der Hamiltonschen partiellen Differentialgleichung (20) sind die geodätischen Nullinien. Die sämtlichen von einem bestimmten Weltpunkt a_s $(s=1,2,3,4)$ ausgehenden geodätischen Nullinien erzeugen eine dreidimensionale Punktmannigfaltigkeit, die die zum Weltpunkt a_s gehörige *Zeitscheide* heißen möge. Die Zeitscheide besitzt in a_s einen Knotenpunkt, dessen Tangentialkegel gerade der zu a_s gehörige Nullkegel ist. Bringen wir die Gleichung der Zeitscheide auf die Gestalt

$$x_4 = \varphi(x_1, x_2, x_3),$$

so ist

$$f = x_4 - \varphi(x_1, x_2, x_3)$$

ein Integral der Hamiltonschen Differentialgleichung (20). Die sämtlichen vom Punkte a_s ausgehenden Zeitlinien verlaufen gänzlich innerhalb desjenigen vierdimensionalen Weltteiles, der die zu a_s gehörige Zeitscheide als Begrenzung hat.

Nach diesen Vorbereitungen wenden wir uns dem Problem der Kausalität in der neuen Physik zu.

Bisher haben wir alle Koordinatensysteme x_s, die aus irgendeinem durch eine willkürliche Transformation hervorgehen, als gleichberechtigt angesehen. Diese Willkür muß eingeschränkt werden, sobald wir die Auffassung zur Geltung bringen wollen, daß zwei auf der nämlichen Zeitlinie gelegene Weltpunkte im Verhältnis von Ursache und Wirkung zueinander stehen können und daß es daher nicht möglich sein soll, solche Weltpunkte auf gleichzeitig zu transformieren. Indem wir x_4 als die *eigentliche* Zeitkoordinate auszeichnen, stellen wir folgende Definitionen auf:

Ein *eigentliches* Raum-Zeit-Koordinatensystem ist ein solches, für welches außer $g < 0$ stets noch die folgenden vier Ungleichungen

$$(21) \qquad g_{11} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0, \qquad \begin{vmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} > 0, \qquad g_{44} < 0$$

erfüllt sind. Eine Transformation, die ein solches Raum-Zeit-Koordinatensystem in ein anderes eigentliches Raum-Zeit-Koordinatensystem überführt, heiße eine *eigentliche* Raum-Zeit-Koordinatentransformation.

Die vier Ungleichungen drücken aus, daß in irgendeinem Weltpunkte a_s der zugehörige Nullkegel den linearen Raum

$$x_4 = a_4$$

ganz außerhalb läßt, die Gerade

$$x_1 = a_1, \qquad x_2 = a_2, \qquad x_3 = a_3$$

dagegen im Inneren enthält; die letztere Gerade ist daher stets eine Zeitlinie.

Es sei nunmehr irgendeine Zeitlinie $x_s = x(p)$ gegeben; wegen

$$G\left(\frac{dx}{dp}\right) < 0$$

folgt dann, daß in einem eigentlichen Raum-Zeit-Koordinatensystem stets

$$\frac{dx_4}{dp} \neq 0$$

sein und folglich längs einer Zeitlinie die eigentliche Zeitkoordinate x_4 stets wachsen bzw. abnehmen muß. Da eine Zeitlinie bei jeder Koordinatentransformation Zeitlinie bleibt, so können zwei Weltpunkte einer Zeitlinie durch eine eigentliche Raum-Zeit-Koordinatentransformation niemals den gleichen Wert der Zeitkoordinate x_4 erhalten d. h. unmöglich auf gleichzeitig transformiert werden.

Andererseits wenn die Punkte einer Kurve eigentlich auf gleichzeitig transformiert werden können, so gilt nach der Transformation für diese Kurve

$$x_4 = \text{konst.}, \quad \text{d. h.} \quad \frac{dx_4}{dp} = 0,$$

mithin

$$G\left(\frac{dx_s}{dp}\right) = \sum_{\mu\nu} g_{\mu\nu} \frac{dx_\mu}{dp} \frac{dx_\nu}{dp} \qquad (\mu, \nu = 1, 2, 3),$$

und hier ist wegen der ersten drei unserer Ungleichungen (21) die rechte Seite positiv; die Kurve charakterisiert sich demnach als eine Strecke.

So sehen wir, daß die dem Kausalitätsprinzip zugrunde liegenden Begriffe von Ursache und Wirkung auch in der neuen Physik zu keinerlei inneren Widersprüchen führen, sobald wir nur stets die Ungleichungen (21) zu unseren Grundgleichungen hinzunehmen, d. h. uns auf den Gebrauch eigentlicher Raum-Zeit-Koordinaten beschränken.

An dieser Stelle sei auf ein späterhin nützliches besonderes Raum-Zeit-Koordinatensystem hingewiesen, welches ich das *Gaußische Koordinatensystem* nennen möchte, weil es die Verallgemeinerung desjenigen geodätischen Polarkoordinatensystems ist, das Gauß in die Flächentheorie eingeführt hat. Es sei in unserer vierdimensionalen Welt irgendein dreidimensionaler Raum gegeben von der Art, daß jede in diesem Raum verlaufende Kurve eine Strecke ist: *ein Streckenraum*, wie ich einen solchen nennen möchte; x_1, x_2, x_3 seien irgendwelche Punktkoordinaten dieses Raumes. Wir konstruieren nun in einem jeden Punkte x_1, x_2, x_3 desselben die zu ihm orthogonale geodätische Linie, die eine Zeitlinie sein wird, und tragen auf derselben x_4 als Eigenzeit auf; dem so erhaltenen Punkte der

vierdimensionalen Welt weisen wir die Koordinatenwerte $x_1 x_2 x_3 x_4$ zu. Für diese Koordinaten wird, wie leicht zu sehen ist,

$$(22) \qquad G(X_s) = \sum_{\mu\nu}^{1,2,3} g_{\mu\nu} X_\mu X_\nu - X_4^2,$$

d. h. das Gaußische Koordinatensystem ist analytisch durch die Gleichungen

$$(23) \qquad g_{14} = 0, \quad g_{24} = 0, \quad g_{34} = 0, \quad g_{44} = -1$$

charakterisiert. Wegen der vorausgesetzten Beschaffenheit des dreidimensionalen Raumes $x_4 = 0$ fällt die rechter Hand in (22) stehende quadratische Form der Variablen X_1, X_2, X_3 notwendig positiv definit aus, d. h. die drei ersten der Ungleichungen (21) sind erfüllt, und da dies auch für die vierte gilt, so erweist sich das Gaußische Koordinatensystem stets als ein *eigentliches* Raum-Zeit-Koordinatensystem.

Wir kehren nun zur Erforschung des Kausalitätsprinzips in der Physik zurück. Als den hauptsächlichen Inhalt desselben sehen wir die Tatsache an, die bisher in jeder physikalischen Theorie galt, daß aus der Kenntnis der physikalischen Größen und ihrer zeitlichen Ableitungen in der Gegenwart allemal die Werte dieser Größen für die Zukunft eindeutig bestimmt werden können: die Gesetze der bisherigen Physik fanden nämlich ausnahmslos ihren Ausdruck in einem System von Differentialgleichungen solcher Art, daß die Anzahl der darin auftretenden Funktionen wesentlich mit der Anzahl der unabhängigen Differentialgleichungen übereinstimmte, und somit bot dann der bekannte Cauchysche Satz über die Existenz von Integralen von Differentialgleichungen unmittelbar den Beweisgrund für jene Tatsache.

Unsere Grundgleichungen (4) und (5) der Physik sind nun keineswegs von der oben charakterisierten Art; vielmehr sind, wie ich gezeigt habe, vier von ihnen eine Folge der übrigen: wir können die elektrodynamischen Gleichungen (5) als Folge der zehn Gravitationsgleichungen (4) ansehen und haben somit für die 14 Potentiale $g_{\mu\nu}$, q_s nur die zehn voneinander wesentlich unabhängigen Gleichungen (4).

Sobald wir an der Forderung der allgemeinen Invarianz für die Grundgleichungen der Physik festhalten, ist der eben genannte Umstand auch wesentlich und notwendig. Gäbe es nämlich für die 14 Potentiale noch weitere von (4) unabhängige invariante Gleichungen, so würde die Einführung eines Gaußischen Koordinatensystems vermöge (23) für die zehn physikalischen Größen

$$g_{\mu\nu} \quad (\mu, \nu = 1, 2, 3), \qquad q_s \quad (s = 1, 2, 3, 4)$$

ein System von Gleichungen liefern, die wiederum voneinander unabhängig wären und, da sie mehr als zehn sind, ein überbestimmtes System bilden würden.

Unter solchen Umständen also, wie sie in der neuen Physik der allgemeinen Relativität zutreffen, ist es keineswegs mehr möglich, aus der Kenntnis der physikalischen Größen in Gegenwart und Vergangenheit eindeutig ihre Werte in der Zukunft zu folgern. Um dies anschaulich an einem Beispiel zu zeigen, seien unsere Grundgleichungen (4) und (5) in dem besonderen Falle integriert, der dem Vorhandensein eines einzigen dauernd ruhenden Elektrons entspricht, so daß sich die 14 Potentiale

$$\begin{aligned} g_{\mu\nu} &= g_{\mu\nu}(x_1, x_2, x_3) \\ q_s &= q_s\ (x_1, x_2, x_3) \end{aligned}$$

als bestimmte Funktionen von x_1, x_2, x_3 ergeben, die von der Zeit x_4 sämtlich unabhängig sind, und überdies so, daß noch die drei ersten Komponenten r^1, r^2, r^3 der Viererdichte verschwinden mögen. Wir wenden sodann auf die Potentiale die folgende Koordinatentransformation an:

$$\begin{aligned} &\begin{cases} x_1 = x_1' & \text{für } x_4' \leqq 0 \\ x_1 = x_1' + e^{-\frac{1}{x_4'^2}} & \text{für } x_4' > 0 \end{cases} \\ &x_2 = x_2' \\ &x_3 = x_3' \\ &x_4 = x_4'; \end{aligned}$$

die transformierten Potentiale $g'_{\mu\nu}, q'_s$ sind für $x'_4 \leqq 0$ die gleichen Funktionen von x'_1, x'_2, x'_3 wie die $g_{\mu\nu}, q_s$ in den ursprünglichen Variablen x_1, x_2, x_3, während die $g'_{\mu\nu}, q'_s$ für $x'_4 > 0$ wesentlich auch von der Zeitkoordinate x'_4 abhängen, d. h. die Potentiale $g'_{\mu\nu}, q'_s$ stellen ein Elektron dar, das bis zur Zeit $x'_4 = 0$ ruht, dann aber sich in seinen Teilen in Bewegung setzt.

Dennoch glaube ich, daß es nur einer schärferen Erfassung der dem Prinzip der allgemeinen Relativität[8]) zugrunde liegenden Idee bedarf, um das Kausalitätsprinzip auch in der neuen Physik aufrechtzuhalten. Dem Wesen des neuen Relativitätsprinzipes entsprechend müssen wir nämlich die Invarianz nicht nur für die allgemeinen Gesetze der Physik verlangen, sondern auch jeder Einzelaussage in der Physik den invarianten Charakter zusprechen, falls sie einen physikalischen Sinn haben soll — im Einklang damit, daß jede physikalische Tatsache letzten Endes durch Lichtuhren, d. h. durch Instrumente von invariantem Charakter fest-

---

[8]) In seiner ursprünglichen, nunmehr verlassenen Theorie hatte A. Einstein (Sitzungsberichte der Akad. zu Berlin 1914, S. 1067) in der Tat, um das Kausalitätsprinzip in der alten Fassung zu retten, gewisse 4 nicht invariante Gleichungen für die $g_{\mu\nu}$ besonders postuliert.

stellbar sein muß. Gerade so wie in der Kurven- und Flächentheorie eine Aussage, für die die Parameterdarstellung der Kurve oder Fläche gewählt ist, für die Kurve oder Fläche selbst keinen geometrischen Sinn hat, wenn nicht die Aussage gegenüber einer beliebigen Transformation der Parameter invariant bleibt oder sich in eine invariante Form bringen läßt, so müssen wir auch in der Physik eine Aussage, die nicht gegenüber jeder beliebigen Transformation des Koordinatensystems invariant bleibt, als *physikalisch sinnlos* bezeichnen. Beispielsweise hat im oben betrachteten Falle des ruhenden Elektrons die Aussage, daß dasselbe etwa zur Zeit $x_4 = 1$ ruhe, physikalisch keinen Sinn, weil diese Aussage nicht invariant ist.

Was nun das Kausalitätsprinzip betrifft, so mögen für die Gegenwart in irgendeinem gegebenen Koordinatensystem die physikalischen Größen und ihre zeitlichen Ableitungen bekannt sein: dann wird eine Aussage nur physikalischen Sinn haben, wenn sie gegenüber allen denjenigen Transformationen invariant ist, bei denen jene als bekannt vorausgesetzten Werte für die Gegenwart unverändert bleiben; ich behaupte, daß die Aussagen dieser Art für die Zukunft sämtlich eindeutig bestimmt sind, d. h. das Kausalitätsprinzip gilt in dieser Fassung:

Aus der Kenntnis der 14 physikalischen Potentiale $g_{\mu\nu}$, q_s in der Gegenwart folgen alle Aussagen über dieselben für die Zukunft notwendig und eindeutig, sofern sie physikalischen Sinn haben.

Um diese Behauptung zu beweisen, benutzen wir das Gaußische Raum-Zeit-Koordinatensystem. Die Einführung von (23) in die Grundgleichungen (4) liefert uns für die 10 Potentiale

$$(24) \qquad g_{\mu\nu} \quad (\mu, \nu = 1, 2, 3), \qquad q_s \quad (s = 1, 2, 3, 4)$$

ein System von ebenso vielen partiellen Differentialgleichungen; wenn wir diese auf Grund der gegebenen Anfangswerte für $x_4 = 0$ integrieren, so finden wir auf eindeutige Weise die Werte von (24) für $x_4 > 0$. Da das Gaußische Koordinatensystem selbst eindeutig festgelegt ist, so sind auch alle auf dieses Koordinatensystem bezogenen Aussagen über jene Potentiale (24) von invariantem Charakter.

Die Formen, in denen physikalisch sinnvolle, d. h. invariante Aussagen mathematisch zum Ausdruck gebracht werden können, sind sehr mannigfaltig.

Erstens. Dies kann mittels eines invarianten Koordinatensystems geschehen. Ebenso wie das vorhin benutzte Gaußische ist zu solchem Zwecke auch das bekannte Riemannsche und desgleichen dasjenige Raum-Zeit-Koordinatensystem verwendbar, in welchem die Elektrizität auf Ruhe und Einheitsdichte transformiert erscheint. Bezeichnet $f(q)$, wie am Schluß

von Teil I die im Hamiltonschen Prinzip auftretende Funktion der Invariante

$$q = \sum_{kl} q_k q_l g^{kl},$$

so ist

$$r^s = 2f'(q) \cdot q^s = 2f'(q) \sum_l g^{sl} q_l$$

die Viererdichte der Elektrizität; sie stellt einen kontravarianten Vektor dar und ist daher für ein Weltgebiet, in dem $f'(q) \neq 0$ ist und das Viererpotential nirgends verschwindet, auf $(0, 0, 0, 1)$ transformierbar. Nach dieser Transformation sind aus den vier Gleichungen

$$\sum_l g^{sl} q_l = 0, \quad (s = 1, 2, 3), \quad \sum_l g^{4l} q_l = \frac{1}{2f'(q)}$$

die vier Komponenten des Viererpotentials q_s durch die $g_{\mu\nu}$ ausdrückbar und jede Beziehung zwischen den $g_{\mu\nu}$ in diesem Koordinatensysteme ist sodann eine invariante Aussage.

Für Partikularlösungen der Grundgleichungen kann es besondere invariante Koordinatensysteme geben; so bilden z. B. im unten behandelten Falle des zentrisch-symmetrischen Gravitationsfeldes r, ϑ, φ, t ein bis auf Drehungen invariantes Koordinatensystem.

Zweitens. Die Aussage, wonach sich ein Koordinatensystem finden läßt, in welchem die 14 Potentiale $g_{\mu\nu}$, q_s für die Zukunft gewisse bestimmte Werte haben oder gewisse bestimmte Beziehungen erfüllen, ist stets eine invariante und daher physikalisch sinnvoll. Der mathematische invariante Ausdruck für eine solche Aussage wird durch Elimination der Koordinaten aus jenen Beziehungen erhalten. Ein Beispiel bietet der oben betrachtete Fall des ruhenden Elektrons: der wesentliche und physikalisch sinnvolle Inhalt des Kausalitätsprinzips drückt sich hier in der Aussage aus, daß das für die Zeit $x_4 \leqq 0$ ruhende Elektron bei geeigneter Wahl des Raum-Zeit-Koordinatensystems auch für die Zukunft $x_4 > 0$ beständig in allen seinen Teilen ruht.

Drittens. Auch ist eine Aussage invariant und hat daher stets physikalischen Sinn, wenn sie für jedes beliebige Koordinatensystem gültig ist, ohne daß dabei die auftretenden Ausdrücke formal invarianten Charakter zu besitzen brauchen.

Nach meinen Ausführungen ist die Physik eine vierdimensionale Pseudogeometrie, deren Maßbestimmung $g_{\mu\nu}$ durch die Grundgleichungen (4) und (5) an die elektromagnetischen Größen, d. h. an die Materie gebunden ist. Mit dieser Erkenntnis wird nun eine alte geometrische Frage zur Lösung reif, die Frage nämlich, ob und in welchem Sinne die Euklidische Geometrie — von der wir aus der Mathematik nur wissen, daß sie

ein logisch widerspruchsfreier Bau ist — auch in der Wirklichkeit Gültigkeit besitzt.

Die alte Physik mit dem absoluten Zeitbegriff übernahm die Sätze der Euklidischen Geometrie und legte sie vorweg einer jeden speziellen physikalischen Theorie zugrunde. Auch Gauß verfuhr nur wenig anders: er konstruierte hypothetisch eine nicht-Euklidische Physik, indem er unter Beibehaltung der absoluten Zeit von den Sätzen der Euklidischen Geometrie nur das Parallelenaxiom fallen ließ; die Messung der Winkel eines Dreieckes mit großen Dimensionen zeigte ihm dann die Ungültigkeit dieser nicht-Euklidischen Physik.

Die neue Physik des Einsteinschen allgemeinen Relativitätsprinzips nimmt gegenüber der Geometrie eine völlig andere Stellung ein. Sie legt weder die Euklidische noch irgendeine andere bestimmte Geometrie vorweg zugrunde, um daraus die eigentlichen physikalischen Gesetze zu deduzieren, sondern die neue Theorie der Physik liefert mit einem Schlage durch ein und dasselbe Hamiltonsche Prinzip die geometrischen und die physikalischen Gesetze, nämlich die Grundgleichungen (4) und (5), welche lehren, wie die Maßbestimmung $g_{\mu\nu}$ — zugleich der mathematische Ausdruck der physkalischen Erscheinung der Gravitation — mit den Werten q_s der elektrodynamischen Potentiale verkettet ist.

Die Euklidische Geometrie ist ein der modernen Physik fremdartiges Ferngesetz: indem die Relativitätstheorie die Euklidische Geometrie als allgemeine Voraussetzung für die Physik ablehnt, lehrt sie vielmehr, daß Geometrie und Physik gleichartigen Charakters sind und als eine Wissenschaft auf gemeinsamer Grundlage ruhen.

Die oben genannte geometrische Frage läuft darauf hinaus, zu untersuchen, ob und unter welchen Voraussetzungen die vierdimensionale Euklidische Pseudogeometrie

$$\begin{array}{llll} g_{11}=1, & g_{22}=1, & g_{33}=1, & g_{44}=-1 \\ g_{\mu\nu}=0 & (\mu \neq \nu) & & \end{array} \tag{25}$$

eine Lösung der Gravitationsgleichungen bzw. die einzige reguläre Lösung derselben ist.

Die Gravitationsgleichungen (8) lauten:

$$[\sqrt{g}\,K]_{\mu\nu}+\frac{\partial\sqrt{g}\,L}{\partial g^{\mu\nu}}=0,$$

wo

$$[\sqrt{g}\,K]_{\mu\nu}=\sqrt{g}\left(K_{\mu\nu}-\frac{1}{2}K g_{\mu\nu}\right)$$

ist. Bei der Einsetzung der Werte (25) wird

$$[\sqrt{g}\,K]_{\mu\nu}=0 \tag{26}$$

und für

$$q_s = 0 \qquad (s = 1, 2, 3, 4)$$

wird

$$\frac{\partial \sqrt{g} L}{\partial g^{\mu\nu}} = 0;$$

d. h. wenn alle Elektrizität entfernt wird, so ist die pseudo-Euklidische Geometrie möglich. Die Frage, ob sie in diesem Falle auch notwendig ist, d. h. ob — bzw. unter gewissen Zusatzbedingungen — die Werte (25) und die durch Transformation der Koordinaten daraus hervorgehenden Werte der $g_{\mu\nu}$ die einzigen regulären Lösungen der Gleichungen (26) sind, ist eine mathematische, hier nicht allgemein zu erörternde Aufgabe. Ich beschränke mich vielmehr darauf, einige besondere diese Aufgabe betreffende Überlegungen anzustellen.

Im Falle der pseudo-Euklidischen Geometrie haben wir

$$g_{\mu\nu} = \gamma_{\mu\nu},$$

worin

$$\gamma_{11} = 1, \quad \gamma_{22} = 1, \quad \gamma_{33} = 1, \quad \gamma_{44} = -1,$$

$$\gamma_{\mu\nu} = 0 \quad (\mu \neq \nu)$$

bedeutet. Für jede dieser pseudo-Euklidischen Geometrie benachbarte Maßbestimmung gilt der Ansatz

$$(27) \qquad g_{\mu\nu} = \gamma_{\mu\nu} + \varepsilon h_{\mu\nu} + \dots,$$

wo ε eine gegen Null konvergierende Größe und $h_{\mu\nu}$ Funktionen der x_s sind. Über die Maßbestimmung (27) mache ich die folgenden zwei Annahmen:

I. Die $h_{\mu\nu}$ mögen von der Variabeln x_4 unabhängig sein.

II. Die $h_{\mu\nu}$ mögen im Unendlichen ein gewisses reguläres Verhalten zeigen.

Soll nun die Maßbestimmung (27) für alle ε die Differentialgleichungen (26) erfüllen, so folgt, daß die $h_{\mu\nu}$ notwendig gewisse lineare homogene partielle Differentialgleichungen zweiter Ordnung erfüllen müssen. Diese Differentialgleichungen lauten, wenn man nach Einstein[*])

$$(28) \qquad h_{\mu\nu} = k_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s k_{ss}, \qquad (k_{\mu\nu} = k_{\nu\mu})$$

$$\delta_{\mu\nu} = 0 \qquad (\mu \neq \nu),$$

$$\delta_{\nu\nu} = 1$$

*) Näherungsweise Integration der Feldgleichungen der Gravitation. Berichte d. Akad. zu Berlin (1916), S. 688.

setzt und zwischen den zehn Funktionen $k_{\mu\nu}$ die vier Relationen

$$\sum_s \frac{\partial k_{\mu s}}{\partial x_s} = 0, \qquad (\mu = 1, 2, 3, 4) \tag{29}$$

annimmt, wie folgt:

$$\Box k_{\mu\nu} = 0, \tag{30}$$

wo zur Abkürzung

$$\Box = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} - \frac{\partial^2}{\partial x_4^2}$$

gesetzt ist.

Die Relationen (29) sind wegen des Ansatzes (28) einschränkende Voraussetzungen für die Funktionen $h_{\mu\nu}$; ich will jedoch zeigen, wie es durch eine geeignete infinitesimale Transformation der Variablen x_1, x_2, x_3, x_4 stets erreicht werden kann, daß für die entsprechenden Funktionen $h'_{\mu\nu}$ nach der Transformation jene einschränkenden Voraussetzungen erfüllt sind.

Zu dem Zwecke bestimme man vier Funktionen $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ der Variablen, die bzw. den Differentialgleichungen

$$\Box \varphi_\mu = \frac{1}{2}\frac{\partial}{\partial x_\mu}\sum_\nu h_{\nu\nu} - \sum_\nu \frac{\partial h_{\mu\nu}}{\partial x_\nu} \tag{31}$$

genügen. Vermöge der infinitesimalen Transformation

$$x_s = x'_s + \varepsilon\varphi_s$$

geht $g_{\mu\nu}$ über in

$$g'_{\mu\nu} = g_{\mu\nu} + \varepsilon\sum_a g_{a\nu}\frac{\partial\varphi_a}{\partial x_\mu} + \varepsilon\sum_a g_{a\mu}\frac{\partial\varphi_a}{\partial x_\nu} + \dots$$

oder wegen (27) in

$$g'_{\mu\nu} = \gamma_{\mu\nu} + \varepsilon h'_{\mu\nu} + \dots,$$

wo

$$h'_{\mu\nu} = h_{\mu\nu} + \frac{\partial\varphi_\nu}{\partial x_\mu} + \frac{\partial\varphi_\mu}{\partial x_\nu}$$

gesetzt ist. Wählen wir nun

$$k_{\mu\nu} = h'_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s h'_{ss},$$

so erfüllen diese Funktionen wegen (31) die Einsteinschen Bedingungen (29) und es wird

$$h'_{\mu\nu} = k_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_s k_{ss} \qquad (k_{\mu\nu} = k_{\nu\mu}).$$

Die Differentialgleichungen (30), die nach den obigen Ausführungen für die gefundenen $k_{\mu\nu}$ gelten müssen, gehen wegen der Annahme I in

$$\frac{\partial^2 k_{\mu\nu}}{\partial x_1^2}+\frac{\partial^2 k_{\mu\nu}}{\partial x_2^2}+\frac{\partial^2 k_{\mu\nu}}{\partial x_3^2}=0$$

über und, da die Annahme II — demgemäß verstanden — zu schließen gestattet, daß die $k_{\mu\nu}$ im Unendlichen sich Konstanten nähern, so folgt, daß dieselben überhaupt Konstante sein müssen, d. h.: Durch Variation der Maßbestimmung der pseudo-Euklidischen Geometrie unter den Annahmen I und II ist es nicht möglich, eine reguläre Maßbestimmung zu erlangen, die nicht ebenfalls pseudo-Euklidisch ist und die doch zugleich einer elektrizitätsfreien Welt entspricht.

Die Integration der partiellen Differentialgleichungen (26) gelingt noch in einem andern Falle, der von Einstein[10]) und Schwarzschild[11]) zuerst behandelt worden ist. Ich gebe im folgenden für diesen Fall einen Weg an, der über die Gravitationspotentiale $g_{\mu\nu}$ im Unendlichen keinerlei Voraussetzungen macht und außerdem auch für meine späteren Untersuchungen Vorteile bietet. Die Annahmen über die $g_{\mu\nu}$ sind folgende:

1. Die Maßbestimmung ist auf ein Gaußisches Koordinatensystem bezogen — nur daß g_{44} noch willkürlich gelassen wird; d. h. es ist

$$g_{14}=0,\qquad g_{24}=0,\qquad g_{34}=0.$$

2. Die $g_{\mu\nu}$ sind von der Zeitkoordinate x_4 unabhängig.

3. Die Gravitation $g_{\mu\nu}$ ist zentrisch symmetrisch in bezug auf den Koordinatenanfangspunkt.

Nach Schwarzschild ist die allgemeinste diesen Annahmen entsprechende Maßbestimmung in räumlichen Polarkoordinaten, wenn

$$\begin{aligned}x_1&=r\cos\vartheta\\x_2&=r\sin\vartheta\cos\varphi\\x_3&=r\sin\vartheta\sin\varphi\\x_4&=t\end{aligned}$$

gesetzt wird, durch den Ausdruck

$$(32)\qquad F(r)\,dr^2+G(r)(d\vartheta^2+\sin^2\vartheta\,d\varphi^2)-H(r)\,dt^2$$

---

[10]) Perihelbewegung des Merkur. Sitzungsber. d. Akad. zu Berlin (1915), S. 831.

[11]) Über das Gravitationsfeld eines Massenpunktes. Sitzungsber. d. Akad. zu Berlin (1916), S. 189.

dargestellt, wo $F(r)$, $G(r)$, $H(r)$ noch willkürliche Funktionen von r sind. Setzen wir

$$r^* = \sqrt{G(r)},$$

so sind wir in gleicher Weise berechtigt r^*, ϑ, φ als räumliche Polarkoordinaten zu deuten. Führen wir in (32) r^* anstatt r ein und lassen dann wieder das Zeichen $*$ weg, so entsteht der Ausdruck

$$(33) \qquad M(r)dr^2 + r^2 d\vartheta^2 + r^2 \sin^2\vartheta\, d\varphi^2 - W(r)dt^2,$$

wo $M(r)$, $W(r)$ die zwei wesentlichen willkürlichen Funktionen von r bedeuten. Die Frage ist, ob und wie diese auf die allgemeinste Weise zu bestimmen sind, damit den Differentialgleichungen (26) Genüge geschieht.

Zu dem Zwecke müssen die bekannten in Teil I angegebenen Ausdrücke $K_{\mu\nu}$, K berechnet werden. Der erste Schritt hierzu ist die Aufstellung der Differentialgleichungen der geodätischen Linien durch Variation des Integrals

$$\int \left(M\left(\frac{dr}{dp}\right)^2 + r^2\left(\frac{d\vartheta}{dp}\right)^2 + r^2 \sin^2\vartheta \left(\frac{d\varphi}{dp}\right)^2 - W\left(\frac{dt}{dp}\right)^2 \right) dp.$$

Wir erhalten als Lagrangesche Gleichungen diese:

$$\frac{d^2 r}{dp^2} + \frac{1}{2}\frac{M'}{M}\left(\frac{dr}{dp}\right)^2 - \frac{r}{M}\left(\frac{d\vartheta}{dp}\right)^2 - \frac{r}{M}\sin^2\vartheta\left(\frac{d\varphi}{dp}\right)^2 + \frac{1}{2}\frac{W'}{M}\left(\frac{dt}{dp}\right)^2 = 0,$$

$$\frac{d^2\vartheta}{dp^2} + \frac{2}{r}\frac{dr}{dp}\frac{d\vartheta}{dp} - \sin\vartheta\cos\vartheta\left(\frac{d\varphi}{dp}\right)^2 = 0,$$

$$\frac{d^2\varphi}{dp^2} + \frac{2}{r}\frac{dr}{dp}\frac{d\varphi}{dp} + 2\,\mathrm{cotg}\,\vartheta\frac{d\vartheta}{dp}\frac{d\varphi}{dp} = 0,$$

$$\frac{d^2 t}{dp^2} + \frac{W'}{W}\frac{dr}{dp}\frac{dt}{dp} = 0;$$

hier und in der folgenden Rechnung bedeutet das Zeichen $'$ die Ableitung nach r. Durch Vergleich mit den allgemeinen Differentialgleichungen der geodätischen Linien:

$$\frac{d^2 x_s}{dp^2} + \sum_{\mu\nu}\left\{ {\mu\nu \atop s} \right\}\frac{dx_\mu}{dp}\frac{dx_\nu}{dp} = 0$$

entnehmen wir für die Klammersymbole $\left\{ {\mu\nu \atop s} \right\}$ die folgenden Werte — wobei die verschwindenden nicht angegeben sind:

$$\left\{ {11 \atop 1} \right\} = \frac{1}{2}\frac{M'}{M}, \quad \left\{ {22 \atop 1} \right\} = -\frac{r}{M}, \quad \left\{ {33 \atop 1} \right\} = -\frac{r}{M}\sin^2\vartheta,$$

$$\left\{ {44 \atop 1} \right\} = \frac{1}{2}\frac{W'}{M}, \quad \left\{ {12 \atop 2} \right\} = \frac{1}{r}, \quad \left\{ {33 \atop 2} \right\} = -\sin\vartheta\cos\vartheta,$$

$$\left\{ {13 \atop 3} \right\} = \frac{1}{r}, \quad \left\{ {23 \atop 3} \right\} = \mathrm{cotg}\,\vartheta, \quad \left\{ {14 \atop 4} \right\} = \frac{1}{2}\frac{W'}{W}.$$

Hiermit bilden wir:

$$
\begin{aligned}
K_{11} = {} & \frac{\partial}{\partial r}\left(\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\} + \left\{ {13 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\right) - \frac{\partial}{\partial r}\left\{ {11 \atop 1} \right\} \\
& + \left\{ {11 \atop 1} \right\}\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\}\left\{ {21 \atop 2} \right\} + \left\{ {13 \atop 3} \right\}\left\{ {31 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\left\{ {41 \atop 4} \right\} \\
& - \left\{ {11 \atop 1} \right\}\left(\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\} + \left\{ {13 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\right) \\
= {} & \frac{1}{2}\frac{W''}{W} - \frac{1}{4}\frac{W'^2}{W^2} - \frac{M'}{rM} - \frac{1}{4}\frac{M'W'}{MW}
\end{aligned}
$$

$$
\begin{aligned}
K_{22} = {} & \frac{\partial}{\partial \vartheta}\left\{ {23 \atop 3} \right\} - \frac{\partial}{\partial r}\left\{ {22 \atop 1} \right\} \\
& + \left\{ {21 \atop 2} \right\}\left\{ {22 \atop 1} \right\} + \left\{ {22 \atop 1} \right\}\left\{ {12 \atop 2} \right\} + \left\{ {23 \atop 3} \right\}\left\{ {32 \atop 3} \right\} \\
& - \left\{ {22 \atop 1} \right\}\left(\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\} + \left\{ {13 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\right) \\
= {} & -1 - \frac{1}{2}\frac{rM'}{M^2} + \frac{1}{M} + \frac{1}{2}\frac{rW'}{MW}
\end{aligned}
$$

$$
\begin{aligned}
K_{33} = {} & -\frac{\partial}{\partial r}\left\{ {33 \atop 1} \right\} - \frac{\partial}{\partial \vartheta}\left\{ {33 \atop 2} \right\} \\
& + \left\{ {31 \atop 3} \right\}\left\{ {33 \atop 1} \right\} + \left\{ {32 \atop 3} \right\}\left\{ {33 \atop 2} \right\} + \left\{ {33 \atop 1} \right\}\left\{ {13 \atop 3} \right\} + \left\{ {33 \atop 2} \right\}\left\{ {23 \atop 3} \right\} \\
& - \left\{ {33 \atop 1} \right\}\left(\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\} + \left\{ {13 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\right) - \left\{ {33 \atop 2} \right\}\left\{ {23 \atop 3} \right\} \\
= {} & \sin^2\vartheta\left(-1 - \frac{1}{2}\frac{rM'}{M^2} + \frac{1}{M} + \frac{1}{2}\frac{rW'}{MW}\right)
\end{aligned}
$$

$$
\begin{aligned}
K_{44} = {} & -\frac{\partial}{\partial r}\left\{ {44 \atop 1} \right\} + \left\{ {41 \atop 4} \right\}\left\{ {44 \atop 1} \right\} + \left\{ {44 \atop 1} \right\}\left\{ {41 \atop 4} \right\} \\
& - \left\{ {44 \atop 1} \right\}\left(\left\{ {11 \atop 1} \right\} + \left\{ {12 \atop 2} \right\} + \left\{ {13 \atop 3} \right\} + \left\{ {14 \atop 4} \right\}\right) \\
= {} & -\frac{1}{2}\frac{W''}{M} + \frac{1}{4}\frac{M'W'}{M^2} + \frac{1}{4}\frac{W'^2}{MW^2} - \frac{W'}{rM}
\end{aligned}
$$

$$
\begin{aligned}
K = \sum_s g^{ss}K_{ss} = {} & \frac{W''}{MW} - \frac{1}{2}\frac{W'^2}{MW^2} - 2\frac{M'}{rM^2} - \frac{1}{2}\frac{M'W'}{M^2W} \\
& - \frac{2}{r^2} + \frac{2}{r^2M} + 2\frac{W'}{rMW}.
\end{aligned}
$$

Wegen

$$\sqrt{g} = \sqrt{MW}\,r^2\sin\vartheta$$

wird

$$K\sqrt{g} = \left\{\left(\frac{r^2W'}{\sqrt{MW}}\right)' - 2\frac{rM'\sqrt{W}}{M^{\frac{3}{2}}} - 2\sqrt{MW} + 2\sqrt{\frac{W}{M}}\right\}\sin\vartheta$$

und, wenn wir

$$M=\frac{r}{r-m},\quad W=w^2\frac{r-m}{r}$$

setzen, wo nunmehr m und w die unbekannten Funktionen von r werden, so erhalten wir schließlich

$$K\sqrt{g}=\left\{\left(\frac{r^2 W'}{\sqrt{MW}}\right)'-2wm'\right\}\sin\vartheta,$$

so daß die Variation des vierfachen Integrals

$$\int\int\int\int K\sqrt{g}\,dr\,d\vartheta\,d\varphi\,dt$$

mit der Variation des einfachen Integrals

$$\int w\,m'\,dr$$

äquivalent ist und zu den Lagrangeschen Gleichungen

$$(34)\qquad \begin{aligned} m'&=0\\ w'&=0\end{aligned}$$

führt. Man überzeugt sich leicht, daß diese Gleichungen in der Tat das Verschwinden sämtlicher $K_{\mu\nu}$ bedingen; sie stellen demnach wesentlich die allgemeinste Lösung der Gleichungen (26) unter den gemachten Annahmen 1., 2., 3., dar. Nehmen wir als Integrale von (34) $m=\alpha$, wo α eine Konstante ist und $w=1$, was offenbar keine wesentliche Einschränkung bedeutet, so ergibt sich aus (33) die gesuchte Maßbestimmung in der von Schwarzschild zuerst gefundenen Gestalt

$$(35)\qquad G(dr,d\vartheta,d\varphi,dt)=\frac{r}{r-\alpha}dr^2+r^2d\vartheta^2+r^2\sin^2\vartheta\,d\varphi^2-\frac{r-\alpha}{r}dt^2.$$

Die Singularität dieser Maßbestimmmung bei $r=0$ fällt nur dann fort, wenn $\alpha=0$ genommen wird, d. h. die Maßbestimmung der pseudo-Euklidischen Geometrie ist bei den Annahmen 1., 2., 3. die einzige reguläre Maßbestimmung, die einer elektrizitätsfreien Welt entspricht.

Für $\alpha\neq 0$ erweisen sich $r=0$ und bei positivem α auch $r=\alpha$ als solche Stellen, an denen die Maßbestimmung nicht regulär ist. Dabei nenne ich eine Maßbestimmung oder ein Gravitationsfeld $g_{\mu\nu}$ an einer Stelle *regulär*, wenn es möglich ist, durch umkehrbar eindeutige Transformation ein solches Koordinatensystem einzuführen, daß für dieses die entsprechenden Funktionen $g'_{\mu\nu}$ an jener Stelle regulär d. h. in ihr und in ihrer Umgebung stetig und beliebig oft differenzierbar sind und eine von Null verschiedene Determinante g' haben.

Obwohl nach meiner Auffassung nur reguläre Lösungen der physikalischen Grundgleichungen die Wirklichkeit unmittelbar darstellen, so

sind doch gerade die Lösungen mit nicht regulären Stellen ein wichtiges mathematisches Mittel zur Annäherung an charakteristische reguläre Lösungen — und in diesem Sinne ist nach dem Vorgange von Einstein und Schwarzschild die für $r=0$ und $r=\alpha$ nicht reguläre Maßbestimmung (35) als Ausdruck der Gravitation einer in der Umgebung des Nullpunktes zentrisch-symmetrisch verteilten Masse anzusehen[12]). Im gleichen Sinne ist auch der Massenpunkt als der Grenzfall einer gewissen Verteilung der Elektrizität um einen Punkt herum aufzufassen, doch sehe ich an dieser Stelle davon ab, die Bewegungsgleichungen desselben aus meinen physikalischen Grundgleichungen abzuleiten. Ähnlich verhält es sich mit der Frage nach den Differentialgleichungen für die Lichtbewegung.

Als Ersatz für die Ableitung aus den Grundgleichungen mögen nach Einstein die folgenden zwei Axiome dienen:

Die Bewegung eines Massenpunktes im Gravitationsfeld wird durch eine geodätische Linie dargestellt, welche Zeitlinie ist.

Die Lichtbewegung im Gravitationsfeld wird durch eine geodätische Nullinie dargestellt[13]).

Da die Weltlinie, die die Bewegung des Massenpunktes darstellt, eine Zeitlinie sein soll, so ist es, wie wir leicht einsehen können, stets möglich, den Massenpunkt durch eigentliche Raum-Zeit-Transformationen auf Ruhe zu bringen, d. h. es gibt eigentliche Raum-Zeit-Koordinatensysteme, in bezug auf die der Massenpunkt beständig ruht.

Die Differentialgleichungen der geodätischen Linien für das zentrische Gravitationsfeld (35) entspringen aus dem Variationsproblem

$$\delta\int\left(\frac{r}{r-\alpha}\left(\frac{dr}{dp}\right)^2+r^2\left(\frac{d\vartheta}{dp}\right)^2+r^2\sin^2\vartheta\left(\frac{d\varphi}{dp}\right)^2-\frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2\right)dp=0,$$

sie lauten nach bekanntem Verfahren:

$$\frac{r}{r-\alpha}\left(\frac{dr}{dp}\right)^2+r^2\left(\frac{d\vartheta}{dp}\right)^2+r^2\sin^2\vartheta\left(\frac{d\varphi}{dp}\right)^2-\frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2=A, \tag{36}$$

$$\frac{d}{dp}\left(r^2\frac{d\vartheta}{dp}\right)-r^2\sin\vartheta\cos\vartheta\left(\frac{d\varphi}{dp}\right)^2=0, \tag{37}$$

$$r^2\sin^2\vartheta\frac{d\varphi}{dp}=B, \tag{38}$$

$$\frac{r-\alpha}{r}\frac{dt}{dp}=C, \tag{39}$$

wo A, B, C Integrationskonstante bedeuten.

---

[12]) Die Stellen $r=\alpha$ nach dem Nullpunkt zu transformieren, wie es Schwarzschild tut, ist meiner Meinung nach nicht zu empfehlen; die Schwarzschildsche Transformation ist überdies nicht die einfachste, die diesen Zweck erreicht.

[13]) Laue hat für den Spezialfall $L=\alpha Q$ gezeigt, wie man diesen Satz aus den elektrodynamischen Gleichungen durch Grenzübergang zur Wellenlänge Null ableiten kann; Phys. Zeitschrift **21** (1920).

Ich beweise zunächst, daß **die Bahnkurven des $r\vartheta\varphi$-Raumes stets in Ebenen liegen, die durch das Gravitationszentrum gehen.**

Zu dem Zwecke eliminieren wir den Parameter p aus den Differentialgleichungen (37) und (38), um so eine Differentialgleichung für ϑ als Funktion von φ zu erhalten. Es ist identisch

$$(40)\quad \frac{d}{dp}\left(r^2\frac{d\vartheta}{dp}\right)=\frac{d}{dp}\left(r^2\frac{d\vartheta}{d\varphi}\cdot\frac{d\varphi}{dp}\right)=\left(2r\frac{dr}{d\varphi}\frac{d\vartheta}{d\varphi}+r^2\frac{d^2\vartheta}{d\varphi^2}\right)\left(\frac{d\varphi}{dp}\right)^2+r^2\frac{d\vartheta}{d\varphi}\frac{d^2\varphi}{dp^2}.$$

Andererseits liefert (38) durch Differentation nach p:

$$\left(2r\frac{dr}{d\varphi}\sin^2\vartheta+2r^2\sin\vartheta\cos\vartheta\frac{d\vartheta}{d\varphi}\right)\left(\frac{d\varphi}{dp}\right)^2+r^2\sin^2\vartheta\frac{d^2\varphi}{dp^2}=0$$

und wenn wir hieraus den Wert von $\frac{d^2\varphi}{dp^2}$ entnehmen und rechter Hand von (40) eintragen, so wird

$$\frac{d}{dp}\left(r^2\frac{d\vartheta}{dp}\right)=\left(\frac{d^2\vartheta}{d\varphi^2}-2\operatorname{ctg}\vartheta\left(\frac{d\vartheta}{d\varphi}\right)^2\right)r^2\left(\frac{d\varphi}{dp}\right)^2.$$

Die Gleichung (37) nimmt damit die Gestalt an:

$$\frac{d^2\vartheta}{d\varphi^2}-2\operatorname{ctg}\vartheta\left(\frac{d\vartheta}{d\varphi}\right)^2=\sin\vartheta\cos\vartheta,$$

eine Differentialgleichung, deren allgemeines Integral

$$\sin\vartheta\cos(\varphi+a)+b\cos\vartheta=0$$

lautet, wo a, b Integrationskonstante bedeuten.

Hiermit ist der gewünschte Nachweis geführt und es genügt daher zur weiteren Diskussion der geodätischen Linien, allein den Wert $\vartheta=\frac{\pi}{2}$ in Betracht zu ziehen. Alsdann vereinfacht sich das Variationsproblem wie folgt

$$\delta\int\left\{\frac{r}{r-\alpha}\left(\frac{dr}{dp}\right)^2+r^2\left(\frac{d\varphi}{dp}\right)^2-\frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2\right\}dp=0,$$

und die drei aus demselben entspringenden Differentialgleichungen erster Ordnung lauten

$$(41)\quad \frac{r}{r-\alpha}\left(\frac{dr}{dp}\right)^2+r^2\left(\frac{d\varphi}{dp}\right)^2-\frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2=A,$$

$$(42)\quad r^2\frac{d\varphi}{dp}=B,$$

$$(43)\quad \frac{r-\alpha}{r}\frac{dt}{dp}=C.$$

Die Lagrangesche Differentialgleichung für r

$$(44)\quad \frac{d}{dp}\left(\frac{2r}{r-\alpha}\frac{dr}{dp}\right)+\frac{\alpha}{(r-\alpha)^2}\left(\frac{dr}{dp}\right)^2-2r\left(\frac{d\varphi}{dp}\right)^2+\frac{\alpha}{r^2}\left(\frac{dt}{dp}\right)^2=0$$

ist notwendig mit den vorigen Gleichungen verkettet und zwar haben wir, wenn die linken Seiten von (41), (42), (43), (44) bzw. mit [1], [2], [3], [4] bezeichnet werden, identisch

$$\frac{d[1]}{dp} - 2\frac{d\varphi}{dp}\frac{d[2]}{dp} + 2\frac{dt}{dp}\frac{d[3]}{dp} = \frac{dr}{dp}[4]. \tag{45}$$

Indem wir $C = 1$ nehmen, was auf eine Multiplikation des Parameters p mit einer Konstanten hinausläuft, und dann aus (41), (42), (43) p und t eliminieren, gelangen wir zu derjenigen Differentialgleichung für $\varrho = \frac{1}{r}$ als Funktion von φ, welche Einstein und Schwarzschild gefunden haben, nämlich:

$$\left(\frac{d\varrho}{d\varphi}\right)^2 = \frac{1+A}{B^2} - \frac{A\alpha}{B^2}\varrho - \varrho^2 + \alpha\varrho^3. \tag{46}$$

Diese Gleichung stellt die Bahnkurve des Massenpunktes in Polarkoordinaten dar; aus ihr folgt in erster Annäherung für $\alpha = 0$ bei $B = \sqrt{\alpha}\, b$, $A = -1 + \alpha a$ die Keplersche Bewegung und die zweite Annäherung führt sodann zu einer der glänzendsten Entdeckungen der Gegenwart: der Berechnung des Vorrückens des Merkurperihels.

Nach dem obigen Axiom soll die Weltlinie für die Bewegung eines Massenpunktes Zeitlinie sein; aus der Definition der Zeitlinie folgt mithin stets $A < 0$.

Wir fragen nun insbesondere, ob der Kreis d. h. $r =$ konst. die Bahnkurve einer Bewegung sein kann. Die Identität (45) zeigt, daß in diesem Falle — wegen $\frac{dr}{dp} = 0$ — die Gleichung (44) keineswegs eine Folge von (41), (42), (43) ist; letztere drei Gleichungen sind daher zur Bestimmung der Bewegung nicht ausreichend; vielmehr sind (42), (43), (44) die notwendig zu erfüllenden Gleichungen. Aus (44) folgt

$$-2r\left(\frac{d\varphi}{dp}\right)^2 + \frac{\alpha}{r^2}\left(\frac{dt}{dp}\right)^2 = 0 \tag{47}$$

oder für die Geschwindigkeit v in der Kreisbahn

$$v^2 = \left(r\frac{d\varphi}{dt}\right)^2 = \frac{\alpha}{2r}. \tag{48}$$

Andererseits ergibt (41) wegen $A < 0$ die Ungleichung

$$r^2\left(\frac{d\varphi}{dp}\right)^2 - \frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2 < 0 \tag{49}$$

oder mit Benutzung von (47)

$$r > \frac{3\alpha}{2}. \tag{50}$$

Wegen (48) folgt hieraus für die Geschwindigkeit des im Kreise sich bewegenden Massenpunktes die Ungleichung[14])

$$v < \frac{1}{\sqrt{3}}. \tag{51}$$

Die Ungleichung (50) gestattet folgende Deutung. Nach (48) ist die Winkelgeschwindigkeit des kreisenden Massenpunktes für $r = r_0$

$$\frac{d\varphi}{dt} = \sqrt{\frac{\alpha}{2r_0^3}}.$$

Wollen wir also statt r, φ die Polarkoordinaten eines um den Nullpunkt mitrotierenden Koordinatensystems einführen, so haben wir nur nötig,

$$\varphi \text{ durch } \varphi + \sqrt{\frac{\alpha}{2r_0^3}}\,t$$

zu ersetzen. Die Maßbestimmung

$$\frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2 - \frac{r-\alpha}{r}dt^2$$

geht durch die betreffende Raum-Zeit-Transformation über in

$$\frac{r}{r-\alpha}dr^2 + r^2 d\varphi^2 + \sqrt{\frac{2\alpha}{r_0^3}}\,r^2 d\varphi\,dt + \left(\frac{\alpha}{2r_0^3}r^2 - \frac{r-\alpha}{r}\right)dt^2.$$

Für $r = r_0$ erhält man hieraus

$$\frac{r_0}{r_0-\alpha}dr^2 + r_0^2 d\varphi^2 + \sqrt{2\alpha r_0}\,d\varphi\,dt + \left(\frac{3\alpha}{2r_0} - 1\right)dt^2$$

und da hier, wegen $r_0 > \frac{3\alpha}{2}$, die Ungleichungen (21) erfüllt sind, so ist für die Umgebung der Bahn des kreisenden Massenpunktes die betrachtete Transformation des Massenpunktes auf Ruhe eine *eigentliche* Raum-Zeit-Transformation.

Andererseits hat auch die in (51) gefundene obere Grenze $\frac{1}{\sqrt{3}}$ für die Geschwindigkeit eines kreisenden Massenpunktes eine einfache Bedeutung. Nach dem Axiom für die Lichtbewegung wird nämlich diese durch eine geodätische Nullinie dargestellt. Setzen wir demnach in (41) $A = 0$, so ergibt sich für die kreisende Lichtbewegung anstatt der Ungleichung (49) die Gleichung

$$r^2\left(\frac{d\varphi}{dp}\right)^2 - \frac{r-\alpha}{r}\left(\frac{dt}{dp}\right)^2 = 0;$$

[14]) Die Angabe von Schwarzschild l. c., wonach sich die Geschwindigkeit des Massenpunktes auf der Kreisbahn bei Verkleinerung des Bahnradius der Grenze $\frac{1}{\sqrt{2}}$ nähert, entspricht der Ungleichung $r \geqq \alpha$ und dürfte nach obigem nicht zutreffend sein.

zusammen mit (47) folgt hieraus für den Radius der Lichtbahn:

$$r = \frac{3\alpha}{2}$$

und für die Geschwindigkeit des kreisenden Lichtes der als obere Grenze in (51) auftretende Wert:

$$v = \frac{1}{\sqrt{3}}.$$

Allgemein erhalten wir für die Lichtbahn aus (46) wegen $A = 0$ die Differentialgleichung

$$(52) \qquad \left(\frac{d\varrho}{d\varphi}\right)^2 = \frac{1}{B^2} - \varrho^2 + \alpha\varrho^3;$$

dieselbe besitzt für $B = \frac{3\sqrt{3}}{2}\alpha$ den Kreis $r = \frac{3\alpha}{2}$ als Poincaréschen „Zykel" — entsprechend dem Umstande, daß alsdann $\varrho - \frac{2}{3\alpha}$ rechts als Doppelfaktor auftritt. In der Tat besitzt in diesem Falle die Differentialgleichung (52) — für die allgemeinere Gleichung (46) gilt Entsprechendes — unendlich viele Integralkurven, die jenem Kreise in Spiralen sich unbegrenzt nähern, wie es die allgemeine Zykeltheorie von Poincaré verlangt.

Betrachten wir einen vom Unendlichen herkommenden Lichtstrahl und nehmen α klein gegenüber seiner kürzesten Entfernung vom Gravitationszentrum, so hat der Lichtstrahl angenähert die Gestalt einer Hyperbel mit Brennpunkt im Zentrum. Daraus ergibt sich auch die Ablenkung, die ein Lichtstrahl durch ein Gravitationszentrum erfährt; dieselbe wird nämlich gleich $\frac{2\alpha}{B}$.

Ein Gegenstück zu der Bewegung im Kreise ist die Bewegung in einer Geraden, die durch das Gravitationszentrum geht. Wir erhalten die Differentialgleichung für diese Bewegung, wenn wir in (44) $\varphi = 0$ setzen und dann aus (43) und (44) p eliminieren; die so entstehende Differentialgleichung für r als Funktion von t lautet:

$$(53) \qquad \frac{d^2r}{dt^2} - \frac{3\alpha}{2r(r-\alpha)}\left(\frac{dr}{dt}\right)^2 + \frac{\alpha(r-\alpha)}{2r^3} = 0$$

mit dem aus (41) folgenden Integral

$$(54) \qquad \left(\frac{dr}{dt}\right)^2 = \left(\frac{r-\alpha}{r}\right)^2 + A\left(\frac{r-\alpha}{r}\right)^3.$$

Nach (53) fällt die Beschleunigung negativ oder positiv aus, d. h. die Gravitation wirkt anziehend oder abstoßend, je nachdem der Absolutwert der Geschwindigkeit

$$\left|\frac{dr}{dt}\right| < \frac{1}{\sqrt{3}}\,\frac{r-\alpha}{r}$$

oder

$$> \frac{1}{\sqrt{3}} \frac{r-\alpha}{r}$$

ausfällt.

Für das Licht ist wegen (54)

$$\left|\frac{dr}{dt}\right| = \frac{r-\alpha}{r};$$

das geradlinig zum Zentrum gerichtete Licht wird in Übereinstimmung mit der letzten Ungleichung stets abgestoßen; seine Geschwindigkeit wächst von 0 bei $r = \alpha$ bis 1 bei $r = \infty$.

Wenn sowohl α wie $\frac{dr}{dt}$ klein sind, geht (53) angenähert in die Newtonsche Gleichung

$$\frac{d^2 r}{dt^2} = -\frac{\alpha}{2} \frac{1}{r^2}$$

über.

(Eingegangen am 29. 12. 1923.)

Über nichtholonome Systeme.

Von

Georg Hamel in Berlin.

Den Anstoß zu dieser Arbeit gaben zwei Noten des Herrn Ivan Tzénoff: „Sur les équations du mouvement des systèmes matériels non holonomes", die 1920 in Liouvilles Journal[1]) und soeben 1924 in diesen Annalen[2]) erschienen sind. Herr Tzénoff leitet Bewegungsgleichungen ab, die einen Mischtyp aus Lagrange und Appell darstellen; die Rechnung kann erheblich vereinfacht werden, so daß man das Resultat sofort einsieht: das soll in § 1 geschehen.

Weiter vergleicht Herr Tzénoff seine Gleichungen mit denen, die Herr Woronetz in seiner Arbeit[3]): „Über die Bewegung eines starren Körpers, der ohne Gleitung auf einer beliebigen Fläche rollt" 1911 in § 5 abgeleitet hat und die inhaltlich mit den von mir in meiner Habilitationsschrift 1903 (auch in der Zeitschrift für Math. und Physik[4]) 1904): „Die Lagrange-Eulerschen Gleichungen der Mechanik" sowie in der Annalenarbeit[5]) „Über die virtuellen Verschiebungen in der Mechanik" gegebenen Bewegungsgleichungen übereinstimmen, nur daß meine Gleichungen erheblich weiter reichen, indem sie die Verwendung beliebiger nichtholonomer Geschwindigkeitsparameter gestatten. In der Form aber sind meine Gleichungen so übersichtlich wie die Lagrangeschen, was man von denen des Herrn Woronetz nicht behaupten kann. Da meine Arbeiten offenbar wenig bekannt sind — Herr Tzénoff nennt sie nicht —, möchte ich in § 2 kurz zeigen, wie die Gleichungen des Herrn Woronetz als Spezialfall in meinen enthalten sind.

---

[1]) Journal de Math. pures et appliquées (8) **3**.

[2]) Math. Annalen **91**.

[3]) Math. Annalen **70**.

[4]) Zeitschr. f. Math. u. Physik **50**.

[5]) Math. Annalen **59**.

Drittens erwähnt Herr Tzénoff am Schlusse seiner Arbeit einen bemerkenswerten Satz des Herrn Woronetz über das Hamiltonsche Prinzip (§ 7 der genannten Arbeit): aber sowohl Herr Woronetz wie auch Herr Tzénoff beweisen die Richtigkeit nur so, daß sie die Übereinstimmung mit ihren Bewegungsgleichungen dartun, während man den Satz auch sofort, fast ohne Rechnung aus dem richtig verstandenen Hamiltonschen Prinzip einsehen kann, wenn man sich der Überlegungen meiner Arbeiten bedient. Das soll in § 3 dieser Note ausgeführt werden. Ich bediene mich der Bezeichnungen meines Lehrbuches über „Elementare Mechanik" (Teubner, Leipzig 1912, 2. Aufl. 1922.)

§ 1.

Die Gleichungen des Herrn Tzénoff.

Es liege ein rheonomes System von einer endlichen Anzahl von Freiheitsgraden vor, d. h. es lasse sich der Ortsvektor nach irgendeinem Punkte des Systems durch

$$\bar{r} = \bar{r}(q_1, q_2, \ldots, q_n, q_{n+1}, \ldots, q_{n+k}; t) \tag{1}$$

darstellen, so daß die kinetische Energie E eine quadratische Funktion der $\dot{q}$ wird, mit Koeffizienten, die noch von den q und von der Zeit t abhängen können.

Ferner mögen noch die folgenden, im allgemeinen nichtholonomen Bedingungen gegeben sein:

$$\dot{q}_{n+m} = \sum_{s=1}^{n} a_{m,s}\,\dot{q}_s + a_m \qquad (m = 1, 2, \ldots, k), \tag{2}$$

denen die Bedingungen für die virtuellen Verschiebungen

$$\delta q_{n+m} = \sum_s a_{m,s}\,\delta q_s \tag{2'}$$

entsprechen. Nach (2) gilt

$$\ddot{q}_{n+m} = \sum_s a_{m,s}\,\ddot{q}_s + \cdots \tag{3}$$

und daher

$$\frac{\partial \ddot{q}_{n+m}}{\partial \ddot{q}_s} = a_{m,s}. \tag{4}$$

Aus (1) folgt in bekannter Weise für die Geschwindigkeit eines Systempunktes

$$\bar{v} = \frac{d\bar{r}}{dt} = \sum_{i=1}^{n} \frac{\partial \bar{r}}{\partial q_i}\,\dot{q}_i + \sum_{s=1}^{k} \frac{\partial \bar{r}}{\partial q_{n+s}}\,\dot{q}_{n+s} + \frac{\partial \bar{r}}{\partial t}$$

und für die Beschleunigung

$$\bar{w} = \frac{d^2\bar{r}}{dt^2} = \sum \frac{\partial \bar{r}}{\partial q_i}\ddot{q}_i + \sum \frac{\partial \bar{r}}{\partial q_{n+s}}\ddot{q}_{n+s} + \dots$$

und daher

(5) $$\frac{\partial \bar{w}}{\partial \ddot{q}_i} = \frac{\partial \bar{r}}{\partial q_i} \qquad (i = 1, 2, \dots, n+k).$$

Weiter folgt aus (1) für die virtuellen Verrückungen

(6) $$\delta\bar{r} = \sum_{i=1}^{n} \frac{\partial \bar{r}}{\partial q_i}\delta q_i + \sum_{s=1}^{k} \frac{\partial \bar{r}}{\partial q_{n+s}}\delta q_{n+s}$$

und daher aus dem Lagrangeschen Prinzip (Vereinigung des Prinzips der virtuellen Arbeiten mit dem von d'Alembert),

$$S\, dm\, \bar{w}\cdot\delta\bar{r} = S\, d\bar{k}\cdot\delta\bar{r},$$

durch Einsetzen von (6)

(7) $$\sum_{i=1}^{n} Q_i\delta q_i + \sum_{s=1}^{k} Q_{n+s}\delta q_{n+s} = \sum_{i=1}^{n} K_i\delta q_i + \sum_{s=1}^{k} K_{n+s}\delta q_{n+s},$$

wo die Q die Lagrangeschen Beschleunigungskomponenten $S\,dm\,\bar{w}\frac{\partial\bar{r}}{\partial q}$ und die K die Kraftkomponenten $S d\bar{k}\frac{\partial \bar{r}}{\partial q}$ sind. (S bedeutet die Summation über das System.)

Setzt man nun (2′) in (7) ein und beachtet, daß die δq_i nun ganz willkürlich sind, so bekommt man die *Bewegungsgleichungen in der Rohform*

(8) $$Q_i + \sum_{m=1}^{k} a_{m,i} Q_{n+m} = K_i + \sum_{m=1}^{k} a_{m,i} K_{n+m} \equiv K_i'.$$

Nun ist bekanntlich

(9) $$Q_i = \frac{d}{dt}\left(\frac{\partial E}{\partial \dot{q}_i}\right) - \frac{\partial E}{\partial q_i},$$

aber nach (5) auch gleich

(10) $$S\,dm\,\bar{w}\frac{\partial \bar{r}}{\partial q_i} = S\,dm\,\bar{w}\frac{\partial \bar{w}}{\partial \ddot{q}_i} = \frac{\partial W}{\partial \ddot{q}_i},$$

wo W die bekannte Appellsche Beschleunigungsfunktion $\frac{1}{2} S\,dm\,\bar{w}^2$ ist.

Nehmen wir in (8) für Q_i den Ausdruck (9), für Q_{n+m} den Ausdruck (10), so bekommen wir

$$\frac{d}{dt}\left(\frac{\partial E}{\partial \dot{q}_i}\right) - \frac{\partial E}{\partial q_i} + \sum_{m=1}^{k} a_{m,i}\frac{\partial W}{\partial \ddot{q}_{n+m}} = K_i',$$

oder wegen (4)

$$\text{(I)} \qquad \frac{d}{dt}\left(\frac{\partial E}{\partial \dot{q}}\right) - \frac{\partial E}{\partial q_i} + \frac{\partial W'}{\partial \ddot{q}_i} = K_i' \qquad (i = 1, 2, \ldots, n),$$

wo der Strich an W' andeuten soll, daß die Differentiation nach den $\ddot{q}_i$ nur insofern vorgenommen werden soll, als sie in den $\ddot{q}_{n+m}$ enhalten sind (nach (3)).

Dies sind die Gleichungen des Herrn Tzénoff in der zweiten Form.

Seine erste Form läßt sich aber auch sofort gewinnen:

Bezeichnet man das E, in das man die $\dot{q}_{n+m}$ nach (2) eingesetzt hat, mit E'', so ist

$$\text{(11)} \qquad \frac{\partial E''}{\partial \dot{q}_i} = \frac{\partial E}{\partial \dot{q}_i} + \sum_m \frac{\partial E}{\partial \dot{q}_{n+m}} \frac{\partial \dot{q}_{n+m}}{\partial \dot{q}_i} = \frac{\partial E}{\partial \dot{q}_i} + \frac{\partial E'}{\partial \dot{q}_i}.$$

wenn auch hier der Strich an E' ähnlich wie oben bei W' andeuten soll, daß nur nach den Variablen differenziert werden soll, die in den $\dot{q}_{n+m}$ vorkommen (nach (2)).

Ebenso ist

$$\text{(12)} \qquad \frac{\partial E''}{\partial q_i} = \frac{\partial E}{\partial q_i} + \sum_m \frac{\partial E}{\partial \dot{q}_{n+m}} \frac{\partial \dot{q}_{n+m}}{\partial q_i} = \frac{\partial E}{\partial q_i} + \frac{\partial E'}{\partial q_i}.$$

Löst man (11) und (12) nach $\frac{\partial E}{\partial \dot{q}_i}$ bzw. nach $\frac{\partial E}{\partial q_i}$ auf und setzt in (I) ein, so erhält man die *erste Form des Herrn Tzénoff:*

$$\text{(II)} \qquad \frac{d}{dt}\left(\frac{\partial E''}{\partial \dot{q}_i}\right) - \frac{\partial E''}{\partial q_i} + \frac{\partial E'}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial E'}{\partial \dot{q}_i}\right) + \frac{\partial W'}{\partial \ddot{q}_i} = K_i'.$$

§ 2.

Die Lagrange-Eulerschen Gleichungen.

In meinen zu Anfang zitierten Arbeiten habe ich folgendes bewiesen (man vergleiche auch die Darstellung in Heuns Lehrbuch der Mechanik **1**, Kinematik, Göschen 1906):

Es sei r eine bloße Funktion von $q_1, q_2, \ldots, q_n$, also E eine bloße Funktion der q und der $\dot{q}$. Man führe neue Geschwindigkeitsgrößen als unabhängige lineare Funktionen der $\dot{q}$ ein,

$$\text{(13)} \qquad \omega_i = \sum_{s=1}^{n} b_{i,s}\, \dot{q}_s \quad \text{oder aufgelöst} \quad \dot{q}_s = \sum_{i=1}^{n} c_{s,i}\, \omega_i .$$

Entsprechend virtuelle Verschiebungen,

$$\text{(14)} \qquad \delta\vartheta_i = \sum_{s=1}^{n} b_{i,s}\, \delta q_s \quad \text{oder aufgelöst} \quad \delta q_s = \sum_{i=1}^{n} c_{s,i}\, \delta\vartheta_i .$$

Aus diesen beiden Definitionen folgen unter der zulässigen, aber nicht gerade notwendigen Annahme

(15) $$\delta d q_i = d \delta q_i,$$

welche mit

$$\delta d \bar{r} = d \delta \bar{r}$$

gleichwertig ist, die *Übergangsgleichungen*

(16) $$d\delta\vartheta_m - \delta d\vartheta_m = \sum_{i,s} \beta_{i,s,m}\, \delta\vartheta_i\, d\vartheta_s,$$

wo $d\vartheta = \omega dt$ und

(17) $$\beta_{i,s,m} = \sum_{h,l} \left(\frac{\partial b_{m,l}}{\partial q_h} - \frac{\partial b_{m,h}}{\partial q_l}\right) c_{l,i}\, c_{h,s}$$

gesetzt ist.

Nützlich, aber nicht notwendig ist es, die $\delta\vartheta_i$ als Konstante aufzufassen und die zweiten Gleichungen (14) als infinitesimale Punkttransformationen, dann geben die zweiten Gleichungen (13) zusammen mit den Übergangsgleichungen (16) die erweiterten Punkttransformationen $\delta d q_s$.

Aus dem Lagrangeschen Prinzip wird dann

(18) $$\sum_i Q_i \delta\vartheta_i = \sum_i K_i \delta\vartheta_i,$$

wo die rechte Seite die virtuelle Arbeit der eingeprägten Kräfte $S d\ddot{k} \delta \bar{r}$ ist, während

(19) $$Q_i = \frac{dJ_i}{dt} - X_i E$$

ist. Dabei bedeutet die Impulsgröße J_i die Ableitung von E nach ω_i, während $X_i E$ die zu $\delta\vartheta_i$ gehörende erweiterte Punkttransformation von E ist. Ausführlich geschrieben ist

(19') $$Q_i = \frac{d}{dt}\left(\frac{\partial E}{\partial \omega_i}\right) + \sum_{s,m} \beta_{i,s,m}\, \omega_s \frac{\partial E}{\partial \omega_m} - \left(\frac{\partial E}{\partial \vartheta_i}\right).$$

Hier ist $\left(\frac{\partial E}{\partial \vartheta_i}\right)$ eine Abkürzung für $\sum_s \frac{\partial E}{\partial q_s} c_{s,i}$. Sind die ϑ_i als wirkliche Koordinaten vorhanden, so ist dieser Ausdruck die Ableitung von E nach ϑ_i, die β sind null und wir haben in (19') den bekannten Lagrangeschen Ausdruck (9) für die Q_i.

Wenn nun die q_i alle unabhängig sind und daher die $\delta\vartheta$ alle willkürlich, *wenn wir also ein holonomes, skleronomes System haben, es aber für gut finden, die nichtholonomen Geschwindigkeitsgrößen ω einzuführen, so lauten die Bewegungsgleichungen nach* (18)

(III)*) $$Q_i = K_i \qquad (i = 1, 2, \ldots, n),$$

*) Die Gleichungen (III) mit (19') schon 1898 bei Volterra: „Sopra una classe di equazioni dinamiche". Atti di Torino **33**. Ferner bei Woronetz, l. c. Gl. (25) § 8.

wo die Q_i nach (19) bzw. (19') zu bestimmen sind und die

$$K_i = \sum_s c_{s,i} S d\bar{k} \frac{\partial \bar{r}}{\partial q_s} = S d\bar{k} \left(\frac{\partial \bar{r}}{\partial \vartheta_i}\right).$$

Wenn aber das System nichtholonom und rheonom ist, so können wir es so einrichten, daß

(20) $$q_n = \vartheta_n = t \quad \text{und also} \quad \dot{q}_n = \omega_n = 1 \quad \text{ist,}$$

und entsprechend

(20') $$\delta q_n = \delta \vartheta_n = 0,$$

daß ferner die nichtholonomen Bedingungsgleichungen die Form annehmen:

(21) $$\omega_{k+h} = 0 \qquad (h = 1, 2, \ldots, n-k-1)$$

und entsprechend die virtuellen Verschiebungen verschwinden:

(21') $$\delta\vartheta_{k+h} = 0 \qquad (h = 1, 2, \ldots, n-k-1),$$

während die k ersten $\delta\vartheta$ frei bleiben.

Infolgedessen bestehen bei den nichtholonomen Bedingungen (21) nur die k ersten der Bewegungsgleichungen (III) und in diesen hat man dann noch (20) und (21) einzusetzen, so daß der Summationsbuchstabe s in (19') nur die Werte 1 bis k und n annimmt. Ferner sieht man: wenn ein ϑ_i eine wirkliche Koordinate ist, sind nach (16) die β mit dem letzten Index i Null und die entsprechenden Glieder können in (19') fortgelassen werden.

Das sind die von mir gegebenen und als Lagrange-Eulersche bezeichneten Gleichungen für nichtholonome Systeme.

Um die Gleichungen des Herrn Woronetz zu bekommen, brauchen wir das Ergebnis nur für den in § 1 geschilderten Fall hinzuschreiben.

Wir setzen

(22) $$\omega_i = \dot{q}_i \quad \text{und} \quad \delta\vartheta_i = \delta q_i \quad \text{für} \quad i = 1, 2, \ldots, n,$$

so daß die β mit einem Schlußindex kleiner oder gleich n null sind.

Weiter setzen wir

(23) $$\omega_{n+m} = \dot{q}_{n+m} - \sum_{s=1}^{k} a_{m,s} \dot{q}_s - a_m \dot{q}_{n+k+1} = 0 \quad \text{für} \quad m = 1, \ldots, k$$

und entsprechend

(23') $$\delta\vartheta_{n+m} = \delta q_{n+m} - \sum_{s=1}^{k} a_{m,s} \delta q_s - a_m \delta q_{n+k+1} = 0 \quad \text{für} \quad m = 1, \ldots, k.$$

Endlich setzen wir

(24) $$\omega_{n+k+1} = \dot{q}_{n+k+1} = 1$$

und

$$(24')\qquad \delta\vartheta_{n+k+1} = \delta q_{n+k+1} = 0,$$

so daß auch die β mit dem letzten Index $n+k+1$ null sind.

Folglich lauten die Bewegungsgleichungen

$$(\mathrm{V})\qquad \frac{dJ_i}{dt} + \sum_{\substack{s=1,2,\ldots,n\\ m=1,2,\ldots,k}} \beta_{i,s,n+m}\,\omega_s J_{n+m} + \sum_{m=1,2,\ldots,k} \beta_{i,\,n+k+1,\,n+m} J_{n+m} - \left(\frac{\partial E}{\partial \vartheta_i}\right) = K_i \qquad (i = 1, 2, \ldots, n).$$

Dabei ist J_i für $i = 1, \ldots, n$ gleich $\frac{\partial E}{\partial \omega_i}$.

Weil aber in allen Fällen *nach* der Differentiation, hier aber auch schon *vor* der Differentiation die Gleichungen (22), (23) und (24) benutzt werden dürfen, ist auch für $i = 1, \ldots, n$

$$J_i = \frac{\partial E''}{\partial \dot q_i}.$$

Dagegen ist

$$J_{n+m} = \frac{\partial E}{\partial \omega_{n+m}} = \frac{\partial E}{\partial \dot q_{n+m}},$$

und hier darf man erst *nach* der Differentiation von den Bedingungsgleichungen (23) und (24) Gebrauch machen.

Rechnet man noch die β nach (23) und (23′) aus, *so hat man in* (V) *genau die Gleichungen des Herrn Woronetz.*

Vergleicht man die verschiedenen Formen von Bewegungsgleichungen miteinander, so ist sicher die Form (19)

$$\frac{dJ}{dt} - XE = K$$

neben der Appellschen

$$\frac{\partial W}{\partial \ddot q} = K$$

die übersichtlichste. Bei welcher die Ausrechnung am einfachsten ist, das wird sehr von den besonderen Umständen abhängen. Oft wird die Rohform (8) unter Benutzung von (9) das einfachste sein. In jedem Falle aber sind die Formen der Herren Appell und Tzénoff deshalb zu beanstanden, weil sie die Funktion W brauchen, während man mit E und den Bedingungsgleichungen auskommt.

§ 3.

Der Satz des Herrn Woronetz.

Aus der Lagrangeschen Zentralgleichung

$$S\,dm\,\bar w\cdot\delta\bar r \equiv \frac{d}{dt} S\,dm\,\bar v\cdot\delta\bar r - \delta E = \delta A$$

ergibt sich sofort durch Integration

$$S\, dm\, \bar{v} \cdot \delta \bar{r} \Big|_{t_1}^{t_2} = \int_{t_1}^{t_2} (\delta E + \delta A)\, dt,$$

und wenn man an den Enden des Zeitintervalls nicht verschiebt, das Hamiltonsche Prinzip:

$$\int (\delta E + \delta A)\, dt = 0. \tag{25}$$

Da die Gültigkeit der Lagrangeschen Zentralgleichung durchaus an die Annahme

$$d\,\delta \bar{r} - \delta\, d\bar{r} = 0$$

geknüpft ist (siehe darüber meine zweite Arbeit), so hat man auch

$$d\,\delta q = \delta\, d q \tag{26}$$

für alle q ohne Ausnahme anzunehmen und hat deshalb, wenn

$$\omega_i = 0$$

eine nichtholonome Bedingung ist, wohl

$$\delta \vartheta_i = 0$$

zu nehmen, aber *nicht*

$$\delta \omega_i = 0,$$

da sonst nach den Übergangsgleichungen die vorstehenden Bedingungen im allgemeinen nicht zu erfüllen wären. Infolgedessen darf man auch in dem Hamiltonschen Prinzip nicht vor Ausführung der Variation von den nichtholonomen Bedingungen Gebrauch machen, sondern erst hinterher.

Nun ist aber, wenn wir wieder von den Annahmen des § 2, insbesondere den Gleichungen (20) bis (21′) Gebrauch machen,

$$\begin{aligned}(\delta E)_{\omega_{k+h}=0,\, \omega_n=1} = {} & \sum_{i=1}^{n} \left(\frac{\partial E}{\partial q_i}\right)_{\omega_{k+h}=0,\, \omega_n=1} \delta q_i + \sum_{i=1}^{k} \left(\frac{\partial E}{\partial \omega_i}\right)_{\omega_{k+h}=0,\, \omega_n=1} \delta \omega_i \\ & + \sum_{l=1}^{n-k-1} \left(\frac{\partial E}{\partial \omega_{k+l}}\right)_{\omega_{k+h}=0,\, \omega_n=1} \delta \omega_{k+l} + \left(\frac{\partial E}{\partial \omega_n}\right)_{\omega_{k+h}=0,\, \omega_n=1} \delta \omega_n .\end{aligned}$$

In den beiden ersten Gliedern darf man auch *vor* Ausführung der Differentiation $\omega_{k+h} = 0$ und $\omega_n = 1$ setzen, so daß die beiden ersten Glieder zusammen $\delta E''$ ergeben, d. h. die Variation von E unter Berücksichtigung der nichtholonomen Bedingungsgleichungen; das letzte Glied fällt fort, da $\delta \omega_n = \frac{\delta d q_n}{dt} = \frac{d \delta t}{dt} = 0$ ist.

Mithin bleibt

$$\delta E = \delta E'' + \sum_{l=1,\dots,n-k-1} J_{k+l}\, \delta \omega_{k+l}$$

und das Hamiltonsche Prinzip (25) nimmt die Form an

$$\text{(VI)} \qquad \int \Big(\delta E'' + \sum_{l=1,2,\ldots,n-k-1} J_{k+l}\,\delta\omega_{k+l} + \delta A\Big)\,dt = 0,$$

wobei (26) für alle q zu beachten ist.

Das ist der Satz des Herrn Woronetz:

„Man darf im Hamiltonschen Prinzip in der kinetischen Energie vor Ausführung der Variation von den nichtholonomen Bedingungen Gebrauch machen, wenn man zur Korrektur dieses Fehlers die Variation der linken Seiten der nichtholonomen Bedingungsgleichungen $\omega_{k+l} = 0$, *jede mit dem zugehörenden Impuls multipliziert, unter dem Integral hinzufügt."*

Berlin, den 3. März 1924.

(Eingegangen am 4. 3. 1924.)

Sur les percussions appliquées aux systèmes matériels.

Von

Ivan Tzénoff in Sofia (Bulgarien).

---

1. Niven a montré (Messenger of Mathematics 4 (1867)) comment on peut appliquer les équations de Lagrange à l'étude des percussions appliquées aux systèmes matériels holonomes. Le même sujet a été traité par Routh (Rigid Dynamics I.). La méthode suivie par ces auteurs n'est pas parfaite, car les équations qu'ils obtiennent contiennent encore les percussions de liaison provenant des liaisons nouvelles introduites au moment de la percussion. Ces équations ne satisfont donc au but poursuivi par Lagrange, qui consiste à obtenir des équations ne contenant aucune force de liaison. M. Appell (Journal de Mathématiques 1896) y arrive, en effectuant un certain changement de paramètres qui servent à déterminer la position du système holonome, au moyen duquel certains nouveaux paramètres (dont le nombre est précisément égal au nombre des liaisons finies qu'on impose subitement au système) deviennent nuls. Après ce changement de paramètres, M. Appell parvient à des formules qui expriment la propriété suivante: Les dérivées partielles de l'énergie cinétique par rapport aux dérivées des paramètres qui ne s'annulent pas au moment du choc, gardent les mêmes valeurs avant et après le choc.

M. M. Beghin et Rousseau (Journal de Mathématiques 1903) en effectuant un changement de paramètres analogue à celui de M. Appell ont montré que la propriété ci-dessus est encore valable pour les systèmes non holonomes. Dans le même journal, M. Appell obtient d'une autre façon les résultats de Beghin et Rousseau.

Nous nous proposons de montrer, comment on peut se servir des équations du mouvement des systèmes établis par nous antérieurement[1]), pour étudier les percussions, appliquées à n'importe quel système, holonome

---

[1]) Voir: 1. Annuaire de l'Université de Sofia 1919; 2. Journal de Math. pures et appliquées 1920; 3. Math. Annalen 91 (1924).

ou non holonome, sans faire aucun changement de paramètres. Nous considérerons premièrement le cas d'une percussion donnée; ensuite le cas où on n'a plus de percussion donnée, mais qu'on impose au système tout d'un coup des liaisons nouvelles, exprimées par des équations finies ou des équations différentielles. Enfin nous déduirons le théorème bien connu de Carnot des équations générales que nous obtiendrons. La méthode que nous donnons est plus simple et plus pratique que celle de M. Appell.

2. *Cas d'une percussion donnée.* Supposons que la position du système non holonome dépende de $k+p$ paramètres $q_1, q_2, \ldots, q_k, \ldots, q_{k+p}$, liés par les p relations différentielles

$$(1) \qquad q'_{k+i} = \sum_{\alpha=1}^{k} a_{i\alpha} q'_\alpha + a_i \qquad (i = 1, 2, \ldots, p)$$

à l'aide desquelles les quantités $q'_{k+1}, q'_{k+2}, \ldots, q'_{k+p}$ sont exprimées en fonction des paramètres et des quantités $q'_1, q'_2, \ldots, q'_k$.

Supposons qu'à partir du moment t_0 une force de percussion donnée commence à agir et dure pendant l'intervalle de temps infiniment court $t_1 - t_0$. Alors le mouvement du système sera troublé et les vitesses de ses différents points éprouveront, pendant le temps $t_1 - t_0$, des variations finies sans que la position du système change sensiblement. Les quantités

$$q'_1, q'_2, \ldots, q'_k, \ldots, q'_{k+p},$$

qui définissent les vitesses, passent dans l'intervalle de temps très court $t_1 - t_0$ des valeurs

$$(q'_1)_0, (q'_2)_0, \ldots, (q'_k)_0, \ldots, (q'_{k+p})_0$$

aux valeurs

$$(q'_1)_1, (q'_2)_1, \ldots, (q'_k)_1, \ldots, (q'_{k+p})_1,$$

tandis que les quantités

$$q_1, q_2, \ldots, q_k, \ldots, q_{k+p},$$

qui définissent la position du système, ne changent pas sensiblement de valeur. Nous ne considérerons que la première approximation qui consiste à négliger l'intervalle $t_1 - t_0$, et supposer que les q' changent brusquement de valeur, tandis que les q ne changent pas.

Désignons par T_0 et S_0 la demi-force vive et la demi-énergie d'accélération du système en ne tenant compte que des liaisons finies imposées au système. D'autre part, désignons par T une quantité analogue à T_0, mais obtenue en tenant compte aussi des relations différentielles (1). Enfin, désignons par T_1 la fonction T_0, considérée comme fonction des quantités $q'_{k+1}, \ldots, q'_{k+p}$ seulement; et par S_1 la fonction S_0, considérée comme fonction des quantités $q''_{k+1}, \ldots, q''_{k+p}$ seulement, en n'oubliant pas

que $q'_{k+1}, \ldots, q'_{k+p}$ et $q''_{k+1}, \ldots, q''_{k+p}$ sont déterminées par les équations (1). Alors l'équation générale de la dynamique, sous la forme que nous lui avons donnée[1]), est

$$\sum_{\alpha=1}^{k}\left(\frac{d}{dt}\frac{\partial T}{\partial q'_\alpha} - \frac{\partial T}{\partial q_\alpha} + \frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha}\right)\delta q_\alpha = \sum_{\alpha=1}^{k} Q_\alpha \delta q_\alpha,$$

ou bien

$$(2)\quad \sum d\left(\frac{\partial T}{\partial q'_\alpha}\right)\delta q_\alpha - \sum \frac{\partial T}{\partial q_\alpha}\delta q_\alpha dt + \sum\left(\frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha}\right)\delta q_\alpha dt = \sum_{\alpha=1}^{k} Q_\alpha \delta q_\alpha dt.$$

On peut appliquer cette équation pour un moment quelconque t de l'intervalle (t_0, t_1), pourvu que tous les δq_α soient arbitraires et indépendants, ce qui signifie, que les déplacements virtuels $(\delta x, \delta y, \delta z)$ des points du système sont compatibles avec toutes les liaisons, finies et différentielles, qui existent à l'instant t. Mais étant donnée la petitesse de l'intervalle $t_1 - t_0$, on peut considérer les liaisons existant au moment t_0 comme invariables pendant l'intervalle $t_1 - t_0$, de façon que les déplacements virtuels qui sont compatibles avec les liaisons à l'instant t_0, le seront encore à l'instant t et inversement. Nous pouvons donc, après avoir choisi pour les δq_α un ensemble de valeurs admissibles au moment t_0, conserver les mêmes valeurs pour toutes les valeurs de t entre t_0 et t_1.

Intégrons maintenant l'équation (2) entre t_0 et t_1, en supposant $t_1 - t_0$ infiniment petit, et que les liaisons, de même que les δq_α, sont invariables pendant cet intervalle de temps; nous obtiendrons

$$(3)\quad \sum \delta q_\alpha\left[\Delta\frac{\partial T}{\partial q'_\alpha} - \int_{t_0}^{t_1}\frac{\partial T}{\partial q_\alpha}dt + \int_{t_0}^{t_1}\left(\frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha}\right)dt\right] = \sum \delta q_\alpha \int_{t_0}^{t_1} Q_\alpha dt,$$

$\Delta\frac{\partial T}{\partial q'_\alpha}$ désignant la variation brusque subie par $\frac{\partial T}{\partial q'_\alpha}$.

La quantité $\frac{\partial T}{\partial q_\alpha}$, étant fonction des paramètres q_α et des quantités q'_α, reste finie; par conséquent l'integrale

$$\int_{t_0}^{t_1}\frac{\partial T}{\partial q_\alpha}dt$$

tend vers zéro avec $t_1 - t_0$. C'est pourquoi nous négligerons ce terme.

L'expression

$$\frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha}$$

est aussi une fonction des q_α et des q'_α, mais ne dépend pas des q''_α. En effet, en nous appuyant sur la relation connue entre les fonctions T_0 et S_0,

$$\frac{\partial S_0}{\partial q''_s} = \frac{d}{dt}\frac{\partial T_0}{\partial q'_s} - \frac{\partial T_0}{\partial q_s} \qquad (s = 1, 2, \ldots, k, \ldots, k+p),$$

et sur les relations (1), nous aurons

$$\frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha} = \sum_{i=1}^{p} \frac{\partial T_0}{\partial q'_{k+i}}\frac{\partial q'_{k+i}}{\partial q_\alpha} - \frac{d}{dt}\sum_{i=1}^{p} \frac{\partial T_0}{\partial q'_{k+i}}\frac{\partial q'_{k+i}}{\partial q'_\alpha} + \sum_{i=1}^{p} \frac{\partial S_0}{\partial q''_{k+i}}\frac{\partial q''_{k+i}}{\partial q''_\alpha}$$

$$= \sum_{i=1}^{p}\left[\frac{\partial T_0}{\partial q'_{k+i}}\left(\frac{\partial q'_{k+i}}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial q'_{k+i}}{\partial q'_\alpha}\right) - \frac{\partial T_0}{\partial q_{k+i}}\frac{\partial q'_{k+i}}{\partial q'_\alpha}\right];$$

d'où l'on voit que la quantité en question ne dépend pas des q''_α; par conséquent elle reste finie et l'intégrale

$$\int_{t_0}^{t_1}\left(\frac{\partial T_1}{\partial q_\alpha} - \frac{d}{dt}\frac{\partial T_1}{\partial q'_\alpha} + \frac{\partial S_1}{\partial q''_\alpha}\right)dt$$

tendra vers zéro avec $t_1 - t_0$ et pourra être négligée.

Si nous posons

$$\int_{t_0}^{t_1} Q_\alpha\, dt = P_\alpha,$$

l'équation (3) prendra la forme

$$\sum_{\alpha=1}^{k} \delta q_\alpha \left(P_\alpha - \Delta \frac{\partial T}{\partial q'_\alpha}\right) = 0. \tag{4}$$

Et comme tous les δq_α sont *indépendants*, nous obtenons le système de k équations

$$\Delta \frac{\partial T}{\partial q'_\alpha} = P_\alpha \qquad (\alpha = 1, 2, \ldots, k). \tag{5}$$

Les premiers membres sont des fonctions linéaires et homogènes des k différences

$$(q'_1)_1 - (q'_1)_0,\ (q'_2)_1 - (q'_2)_0,\ \ldots,\ (q'_k)_1 - (q'_k)_0.$$

Comme les $(q'_1)_0, (q'_2)_0, \ldots, (q'_k)_0$ sont connues, on en déduit les $(q'_1)_1, (q'_2)_1, \ldots, (q'_k)_1$. Enfin des relations (1), on tire les $(q'_{k+1})_1, \ldots, (q'_{k+p})_1$, ce qui achève la détermination des quantités $q'_1, q'_2, \ldots, q'_k, \ldots, q'_{k+p}$ après la percussion, et par conséquent la détermination des vitesses après le choc.

Remarques: 1) Pendant le temps $t_1 - t_0$ infiniment court que dure la percussion, prennent naissance *des percussions de liaison*, dues aux liaisons existant pendant la durée de la percussion. Mais ces percussions n'interviennent pas dans l'équation générale de la dynamique, car, les déplacements virtuels du système étant compatibles avec les liaisons

existant pendant le temps $t_1 - t_0$, la somme des travaux des percussions de liaison est nulle, en supposant bien entendu que les liaisons sont réalisées sans frottement. On vérifie cette propriété de la même manière que dans le cas de forces de liaison ordinaires. Cela tient au fait que les percussions de liaison sont dues aux forces de liaison, agissant pendant la durée $t_1 - t_0$, et que les corps du système ne se sont pas sensiblement déplacés pendant la durée de la percussion. Par exemple, si nous avons deux corps S et S' en contact sans frottement, les forces de liaison sont normales au plan tangent commun et sont égales et directement opposées. D'où il s'ensuit que les percussions résultant de cette liaison sont aussi normales au plan tangent commun et sont égales et directement opposées. En effet, en désignant par α, β, γ les cosinus-directeurs de la normale commune, et par N la réaction de S' sur S, les projections de N seront $N\alpha$, $N\beta$, $N\gamma$. Pendant le temps $t_1 - t_0$ N devient très grand et donne naissance à une percussion de liaison, qui a pour projections

$$\int_{t_0}^{t_1} N\alpha\, dt, \quad \int_{t_0}^{t_1} N\beta\, dt, \quad \int_{t_0}^{t_1} N\gamma\, dt.$$

Mais comme le corps ne se déplace presque pas pendant la durée $t_1 - t_0$, nous pouvons considérer α, β, γ comme indépendants de t et écrire ces projections de la sorte

$$\alpha\int_{t_0}^{t_1} N dt, \quad \beta\int_{t_0}^{t_1} N dt, \quad \gamma\int_{t_0}^{t_1} N dt.$$

La percussion de liaison de S' sur S est le vecteur

$$P = \int_{t_0}^{t_1} N dt,$$

normal au plan tangent commun. De même, la percussion de liaison de S sur S' est un vecteur P', égal et directement opposé à P, car ses projections se déduisent des précédentes en changeant les signes. Alors, pour un déplacement compatible avec la liaison, la somme des travaux des percussions P et P' est nulle.

2) Les raisonnements précédents restent encore valables pour le cas particulier, où le système donné est holonome, car il est évident que l'équation (4) est la même, que le système soit holonome ou non. Dans ce cas, les équations (5) suffisent pour déterminer l'état des vitesses après la percussion, parce que dans ce cas le nombre de ces équations est exactement égal au nombre des paramètres q_α du système holonome.

3. *Cas où il n'existe pas de percussion donnée, mais qu'on introduit subitement de nouvelles liaisons persistantes.* Si on introduit dans le système, pendant la durée très courte que dure la percussion, de

nouvelles liaisons persistantes, alors pour en tenir compte, il ne faudra considérer, parmi les déplacements compatibles avec les liaisons existant antérieurement que ceux qui le sont aussi avec les liaisons nouvelles. Par conséquent, les nouvelles liaisons n'ont pour effet que de restreindre le champ des déplacements virtuels. Il est rare que l'introduction des nouvelles liaisons coïncide avec le moment où agit la percussion donnée.

Mais ce qui arrive souvent c'est que, sans avoir fait agir sur le système une percussion donnée, on introduit subitement à l'instant t_0 de nouvelles liaisons persistantes; alors pendant l'intervalle de temps très court $t_1 - t_0$, les vitesses des differents points du système éprouvent des variations finies sans que le système change de position. Alors dans le système prennent naissance des percussions dues aux nouvelles forces de liaison, c. à d. des percussions de liaison, qu'on doit laisser de côté, comme nous l'avons expliqué plus haut.

Voyons maintenant quelles sont les équations qui déterminent les variations des vitesses.

Supposons maintenant que les nouvelles liaisons sont au nombre de r $(r<k)$. Elles peuvent être finies ou différentielles et représentent certaines relations qui existent entre les paramètres q_a et leurs dérivées q'_a $(a=1, 2, \ldots, k)$. Supposons que les nouvelles liaisons soient données par les équations

$$q'_{n+j} = \sum_{f=1}^{n} b_{jf} q'_f + b_j \qquad (j=1, 2, \ldots, r), \quad (n=k-r) \tag{6}$$

d'où l'on tire

$$\delta q_{n+j} = \sum_{f=1}^{n} b_{jf} \delta q_f \qquad (j=1, 2, \ldots, r). \tag{7}$$

Comme il n'y a pas de percussions données $P_a = 0$ et l'équation (4) prend la forme

$$\sum_{f=1}^{n} \delta q_f \varDelta \frac{\partial T}{\partial q'_f} + \sum_{j=1}^{r} \delta q_{n+j} \varDelta \frac{\partial T}{\partial q'_{n+j}} = 0,$$

ou, en tenant compte des équations (7), et que d'après (6)

$$\frac{\partial q'_{n+j}}{\partial q'_f} = b_{jf} \qquad (j=1, 2, \ldots, r),$$

nous obtenons

$$\sum_{f=1}^{n} \left(\varDelta \frac{\partial T}{\partial q'_f} + \sum_{j=1}^{r} \varDelta \frac{\partial T}{\partial q'_{n+j}} \cdot \frac{\partial q'_{n+j}}{\partial q'_f} \right) \delta q_f = 0.$$

La quantité T (l'énergie cinétique du système) est une fonction des coordonnées $q_1, q_2, \ldots, q_k, \ldots, q_{k+p}$ du système et des quantités

$$q'_1, q'_2, \ldots, q'_n, q'_{n+1}, \ldots, q'_{n+r=k},$$

les $q'_{n+1}, \ldots, q'_{n+r}$ étant déterminées par les équations (6). Alors, si nous désignons par $\overline{T}$ la fonction T, considérée comme une fonction des quantités ci-dessus, mais en tenant compte des relations (6), la dernière équation obtenue prend la forme

$$\sum_{f=1}^{n} \Delta \frac{\partial \overline{T}}{\partial q'_f} \delta q_f = 0.$$

Comme les δq_f $(f = 1, 2, \ldots, n)$ sont arbitraires, cette équation nous donne les n équations

$$\Delta \frac{\partial \overline{T}}{\partial q'_f} = 0 \qquad (f = 1, 2, \ldots, n), \tag{8}$$

ou bien

$$\left(\frac{\partial \overline{T}}{\partial q'_f}\right)_1 = \left(\frac{\partial \overline{T}}{\partial q'_f}\right)_0 \qquad (f = 1, 2, \ldots, n); \tag{8'}$$

ces équations sont linéaires et homogènes par rapport aux k différences

$$(q'_1)_1 - (q'_1)_0, \ldots, (q'_n)_1 - (q'_n)_0, \ldots, (q'_k)_1 - (q'_k)_0.$$

A l'instant t_0, où la percussion commence, les quantités

$$(q'_1)_0, \ldots, (q'_n)_0, \ldots, (q'_k)_0$$

sont indépendantes et ont des valeurs données; à l'instant t_1, où la percussion finit, les quantités

$$(q'_1)_1, (q'_2)_1, \ldots, (q'_n)_1$$

sont indépendantes et leurs valeurs sont déterminées par des équations (8); tandis que les autres quantités

$$(q'_{n+1})_1, (q'_{n+2})_1, \ldots, (q'_k)_1$$

sont déterminées par les relations différentielles (6), prises à l'instant t_1, c. à. d. par les équations

$$(q'_{n+j})_1 = \sum_{f=1}^{n} b_{jf} (q'_f)_1 \qquad (j = 1, 2, \ldots, r). \tag{9}$$

Donc, les équations (8) et (9) au nombre de $n + r = k$ déterminent complètement les

$$(q'_1)_1, (q'_2)_1, \ldots, (q'_k)_1$$

et par conséquent les équations (1), (8) et (9) déterminent l'état des vitesses après les percussions.

Les équations (8) expriment la propriété suivante:

Les dérivées partielles de la fonction T, énergie cinétique du système, où les quantités $q'_1, q'_2, \ldots, q'_n$ sont indépendantes et les $q'_{n+1}, q'_{n+2}, \ldots, q'_k$

sont déterminées par les liaisons subitement introduites, par rapport aux variables indépendantes, ont les mêmes valeurs avant et après le choc.

Remarques. 1) Les raisonnements précedents restent en vigueur aussi pour un système holonome, dans lequel on introduit brusquement de nouvelles liaisons, exprimées par des liaisons finies ou différentielles, puisque nous avons vu que l'équation (4), dont nous avons tiré notre théorème, est la même pour les systèmes holonomes et non holonomes.

2) Il faut remarquer qu'il y a une différence entre les liaisons exprimées par l'équation (1), et celles exprimées par les équations (6). Les deux sortes de liaisons sont persistantes, c. à. d. elles existent aussi après le choc, comme pendant le choc, mais les liaisons (1) existent avant le choc, tandis que les liaisons (6) n'existaient pas avant le choc. Par conséquent, nous n'avons pas le droit de substituer dès le début dans T à $q'_{n+1}, q'_{n+2}, \ldots, q'_{n+r}$ leurs valeurs tirées de (6) et calculer alors les dérivées partielles par rapport à $q'_1, q'_2, \ldots, q'_n$, les seules quantités dont elle dépend; pour faire ce calcul il faut faire comme nous l'avons expliqué plus haut, c. à. d. d'après la règle de dérivation des fonctions composées.

3) Si les nouvelles liaisons subitement introduites, exprimées par les équations (6), existent seulement pendant le choc, mais non pas après, c. à. d. si les nouvelles liaisons *ne sont pas persistantes*, les équations (8) au nombre de n, servent à déterminer les $n+r=k$ inconnues

$$(q'_1)_1, (q'_2)_1, \ldots, (q'_n)_1, \ldots, (q'_k)_1;$$

les équations (9) dans ce cas n'existent pas. Pour pouvoir déterminer toutes les inconnues, il nous faudra encore r autres équations. Pour les obtenir, il faut faire des hypothèses particulières sur les phénomènes qui se déroulent après le choc subi par le système. Considérons par exemple le choc direct de deux sphères, ayant des rayons R_1 et R_2, des masses m_1 et m_2, dont les centres se déplacent sur un axe fixe ox. Supposons que les deux sphères effectuent des mouvements de translation. Désignons par x_1 et x_2 les abscisses des centres des deux sphères; la position du système dépend des deux paramètres x_1 et x_2. Au moment du choc une nouvelle liaison s'introduit: elle s'exprime par l'équation finie

$$x_2 - x_1 - R_1 - R_2 = 0,$$

qui exprime que la distance des centres est égale à la somme des rayons.

La fonction T sera

$$T = \frac{1}{2}(m_1 x_1'^2 + m_2 x_2'^2),$$

et l'équation (6) sera

$$x_2' = x_1'. \tag{10}$$

Alors d'après l'équation (8) nous aurons

$$\left(\frac{\partial \overline{T}}{\partial x_1'}\right)_1 = \left(\frac{\partial \overline{T}}{\partial x_1'}\right)_0,$$

c. à. d.

(11) $$m_1(x_1')_1 + m_2(x_2')_1 = m_1(x_1')_0 + m_2(x_2')_0;$$

cette équation exprime que la somme des projections sur l'axe ox des quantités de mouvement ne change pas.

Si la liaison (10) est persistante, c. à. d. si les sphères sont parfaitement molles, l'équation (10) subsistera après le choc et nous donnera

(12) $$(x_2')_1 = (x_1')_1;$$

équation qui remplace les équations (9). Alors les équations (11) et (12) suffisent pour déterminer les inconnues $(x_1')_1$ et $(x_2')_1$.

Si la liaison subitement introduite (10) ne persiste pas après le choc, c. à. d. n'est pas persistante, alors l'équation (12) n'existe pas et l'équation (11) ne suffit pas pour la détermination des inconnues. Ce cas se présente quand les sphères sont parfaitement élastiques, c. à. d. quand il n'y a aucune perte de force vive. Dans ce cas nous aurons

(13) $$m_1(x_1')_1^2 + m_2(x_2')_1^2 = m_1(x_1')_0^2 + m_2(x_2')_0^2.$$

Des équations (11) et (13) nous pouvons déterminer $(x_1')_1$ et $(x_2')_1$.

Nous allons maintenant traiter quelques problèmes, où les liaisons introduites sont persistantes.

4. Exemples. 1) *Pendule balistique.* Supposons que le projectile se déplace dans un plan vertical d'un mouvement de translation; prenons pour origine des coordonnées la trace de l'axe de suspension sur ce plan. Désignons par r et α les coordonnées polaires du projectile, par θ, l'angle de déviation du pendule de la verticale. La position du système dépend de trois paramètres r, α et θ. Après le choc, le pendule composé formé par le projectile et le pendule balistique tourne autour de l'axe de suspension. Par conséquent, r reste constant $(= a)$, α devient égal à θ. Les liaisons introduites sont:

(14) $$r = a, \quad \alpha = \theta$$

ou bien

(14') $$r' = 0, \quad \alpha' = \theta'.$$

L'énergie cinétique du système est

$$T = \frac{1}{2}[I\theta'^2 + m(r'^2 + r^2\alpha'^2)].$$

I désignant le moment d'inertie du pendule par rapport à l'axe de suspension. Mais comme on a introduit subitement deux liaisons nouvelles, nous n'aurons à écrire qu'une seule équation, notamment celle qui se rapporte à θ

$$\left(\frac{\partial T}{\partial \theta'}\right)_1 = \left(\frac{\partial T}{\partial \theta'}\right)_0,$$

ou

$$I(\theta')_1 + mr^2(\alpha')_1 = I(\theta')_0 + mr^2(\alpha')_0,$$

mais d'après (14′) et (14) nous avons $(\alpha')_1 = (\theta')$, $r = a$. Donc

$$(I + ma^2)(\theta')_1 = ma^2(\alpha')_0,$$

$(\theta')_0$ étant nul puisque le pendule part du repos. L'équation obtenue nous donnera la vitesse angulaire finale $(\theta')_1$ du pendule, puisque $a(\alpha')_0$ est connu; ceci représente en effet la projection de la vitesse du projectile m sur la perpendiculaire au rayon om au moment du choc.

2) Un disque circulaire homogène de masse M et de rayon R se déplace dans un plan vertical xoy; à l'instant t_0 il heurte un axe fixe ox et depuis lors ne peut que rouler sur cet axe. Déterminons la vitesse après le choc.

La position du système avant le choc dépend de trois paramètres: les coordonnées ξ et η du centre du disque et l'angle θ, dont il a tourné dans le sens négatif de oy vers ox. Au moment du choc, deux nouvelles liaisons s'introduisent:

1) le disque touche l'axe ox, donc

$$\eta = R;$$

2) il roule sur l'axe ox, donc

$$\xi = R\theta,$$

en choisissant convenablement la position initiale. Nous aurons

$$\eta' = 0, \quad \xi' = R\theta'.$$

L'énergie cinétique sera

$$T = \frac{M}{2}(k^2\theta'^2 + \xi'^2 + \eta'^2),$$

Mk^2 désignant le moment d'inertie du disque par rapport à son centre. L'équation cherchée est

$$\left(\frac{\partial T}{\partial \theta'}\right)_1 = \left(\frac{\partial T}{\partial \theta'}\right)_0,$$

ou

$$k^2(\theta')_1 + R(\xi')_1 = k^2(\theta')_0 + R(\xi')_0,$$

mais $(\xi')_1 = R(\theta')_1$; alors

$$(\theta')_1 = \frac{k^2(\theta')_0 + R(\xi')_0}{k^2 + R^2}$$

ou

$$(\xi')_1 = R\,\frac{k^2(\theta')_0 + R(\xi')_0}{k^2 + R^2}.$$

Cette formule donne la vitesse finale du centre dans le mouvement de roulement.

3) Une sphère homogène de rayon a se meut dans l'espace; à l'instant t_0 elle heurte un plan horizontal, après quoi elle ne peut que rouler sans glisser sur celui-ci. Déterminer sa vitesse après le choc.

Choisissons dans le plan deux axes fixes $o\xi$, $o\eta$ et un troisième axe $o\zeta$ perpendiculaire au plan. Soient ξ, η, ζ les coordonnées du centre de gravité G de la sphère. Menons par G trois axes $G x_1 y_1 z_1$, parallèles aux axes $o\xi\eta\zeta$ et désignons par p_1, q_1, r_1 les projections de la rotation instantanée de la sphère sur ces axes. En prenant la masse de la sphère égale à un, soit $A = Mk^2 = k^2$ son moment d'inertie par rapport à un quelconque de ses diamètres.

L'énergie cinétique de la sphère est

$$T = \frac{1}{2}\sum m v^2 = \frac{1}{2}[\xi'^2 + \eta'^2 + \zeta'^2 + k^2(p_1^2 + q_1^2 + r_1^2)].$$

En désignant par θ, φ, ψ les angles d'Euler avec les axes $G x_1 y_1 z_1$, d'un système d'axes $G x y z$ liés à la sphère nous avons les formules:

$$(15)\qquad \begin{cases} p_1 = \theta' \cos\psi + \varphi' \sin\theta \sin\psi, \\ q_1 = \theta' \sin\psi - \varphi' \sin\theta \cos\psi, \\ r_1 = \psi' + \varphi' \cos\theta \end{cases}$$

et la fonction T devient

$$(16)\qquad 2T = \xi'^2 + \eta'^2 + \zeta'^2 + k^2(\theta'^2 + \varphi'^2 + \psi'^2 + 2\psi'\varphi'\cos\theta).$$

Au moment, où le choc avec le plan horizontal se produit, il s'introduit subitement une nouvelle liaison — la sphère roule sans glisser sur le plan. On exprimera cette condition en écrivant que le point de contact de la sphère avec le plan a une vitesse nulle. Comme ce point a pour coordonnées $0, 0, -a$ par rapport aux axes $G x_1 y_1 z_1$, nous aurons évidemment

$$\xi' - a q_1 = 0, \qquad \eta' + a p_1 = 0, \qquad \zeta' = 0.$$

ou bien d'après les équations (15)

$$(17)\qquad \begin{cases} \xi' = a(\theta' \sin\psi - \varphi' \sin\theta\cos\psi), \\ \eta' = -a(\theta'\cos\psi + \varphi'\sin\theta\sin\psi), \\ \zeta' = 0; \end{cases}$$

équations qui remplacent les équations (6).

Les équations (8) dans ce cas sont les suivantes:

$$\left(\frac{\partial \bar{T}}{\partial \theta'}\right)_1 - \left(\frac{\partial \bar{T}}{\partial \theta'}\right)_0 = 0,$$

$$\left(\frac{\partial \bar{T}}{\partial \psi'}\right)_1 - \left(\frac{\partial \bar{T}}{\partial \psi'}\right)_0 = 0,$$

$$\left(\frac{\partial \bar{T}}{\partial \varphi'}\right)_1 = \left(\frac{\partial \bar{T}}{\partial \varphi'}\right)_0 = 0.$$

L'équation pour θ' est

$$k^2\theta_1' + \xi_1' a\sin\psi - \eta_1' a\cos\psi = k^2\theta_0' + \xi_0' a\sin\psi - \eta_0' a\cos\psi;$$

mais comme

$$(18)\qquad \begin{cases} \xi_1' = a(\theta_1' \sin\psi - \varphi_1' \sin\theta\cos\psi), \\ \eta_1' = -a(\theta_1'\cos\psi + \varphi_1'\sin\theta\sin\psi), \\ \zeta_1' = 0, \end{cases}$$

on a

$$(19)\qquad \theta_1' = \frac{k^2\theta_0' + \xi_0' a\sin\psi - \eta_0' a\cos\psi}{k^2 + a^2}.$$

L'équation pour ψ' donne

$$(20)\qquad \psi_1' + \varphi_1'\cos\theta = \psi_0' + \varphi_0'\cos\theta,$$

ce qui montre que la projection de la rotation instantanée de la sphère sur l'axe Gz_1 ne change pas avant et après le choc.

Enfin l'équation pour φ' sera

$$k^2(\varphi_1' + \psi_1'\cos\theta) - \xi_1' a\sin\theta\cos\psi - \eta' a\sin\theta\sin\psi$$
$$= k^2(\varphi_0' + \psi_0'\cos\theta) - \xi_0' a\sin\theta\cos\psi - \eta_0' a\sin\theta\sin\psi,$$

ou bien d'après (18)

$$(21)\qquad k^2(\varphi_1' + \psi_1'\cos\theta) + \varphi_1' a^2\sin^2\theta$$
$$= k^2(\varphi_0' + \psi_0'\cos\theta) - \xi_0' a\sin\theta\cos\psi - \eta_0' a\sin\theta\sin\psi.$$

Les équations (20) et (21) résolues par rapport à φ_1' et ψ_1' donnent

$$(22)\qquad \varphi_1' = \frac{k^2\sin\theta\,\varphi_0' - \xi_0' a\cos\psi - \eta_0' a\sin\psi}{(k^2 + a^2)\sin\theta},$$

$$(23)\qquad \psi_1' = \psi_0' + \varphi_0' \cos\theta - \cos\theta \frac{k^2 \sin\theta \cdot \varphi_0' - \xi_0' a \cos\psi - \eta_0' a \sin\psi}{(k^2 + a^2)\sin\theta}.$$

En tenant compte des équations (19) et (20), nous tirons des équations (15) et (18) les valeurs pour ξ_1' et η_1':

$$(24)\qquad \begin{cases} \xi_1' = \dfrac{k^2 a (q_1)_0 + \xi_0' a^2}{k^2 + a^2} \\ \eta_1' = \dfrac{-k^2 a (p_1)_0 + \eta_0' a^2}{k^2 + a^2} \\ \zeta_1' = 0. \end{cases}$$

Les équations (24), (19), (22), (23) déterminent la vitesse de la sphère après le choc.

Les projections de la rotation instantanée de la sphère sur les axes $G x_1 y_1 z_1$ après le choc seront

$$(25)\qquad \begin{cases} (p_1)_1 = \dfrac{k^2 (p_1)_0 + \eta_0' a}{k^2 + a^2} \\ (q_1)_1 = \dfrac{k^2 (q_1)_0 + \xi_0' a}{k^2 + a^2} \\ (r_1)_1 = (r_1)_0 . \end{cases}$$

5. *Théorème de Carnot.* L'énoncé de ce théorème bien connu est le suivant: ***Si dans un système soumis à des liaisons sans frottement, on introduit de nouvelles liaisons persistantes, la force vive perdue est égale à la force vive due aux vitesses perdues.***

Nous allons le déduire des équations générales (8').

Supposons toujours que la position du système dépende de $k + p$ paramètres: $q_1, q_2, \ldots, q_n, \ldots, q_k, \ldots, q_{k+p}$, entre lesquels il existe p relations différentielles, ne contenant pas le temps t, de la forme

$$(26)\qquad q_{k+i}' = \sum_{\alpha=1}^{k} a_{i\alpha} q_\alpha' \qquad (i = 1, 2, \ldots, p).$$

Supposons qu'à l'instant t_0 on introduise r nouvelles liaisons persistantes de la forme

$$(27)\qquad q_{n+j}' = \sum_{f=1}^{n} b_{if} q_f', \qquad (j = 1, 2, \ldots, r).$$

r étant $< k$ et $k - r = n$.

Les équations (8') sont

$$\left(\frac{\partial \overline{T}}{\partial q_f'}\right)_1 - \left(\frac{\partial \overline{T}}{\partial q_f'}\right)_0 = 0 \qquad (f = 1, 2, \ldots, n).$$

Les coordonnées d'un point quelconque du système sont:

$$x_0 = f(q_1, q_2, \ldots, q_n, \ldots, q_k, \ldots, q_{k+p}),$$
$$y_0 = \psi(q_1, q_2, \ldots, q_n, \ldots, q_k, \ldots, q_{k+p}),$$
$$z_0 = \omega(q_1, q_2, \ldots, q_n, \ldots, q_k, \ldots, q_{k+p}).$$

Les projections de sa vitesse x_0', y_0', z_0' seront:

$$x_0' = \sum_{\alpha=1}^{k} \frac{\partial x_0}{\partial q_\alpha} q_\alpha' + \sum_{i=1}^{p} \frac{\partial x_0}{\partial q_{k+i}} q_{k+i}', \qquad y_0 = \ldots, \qquad z_0 \ldots,$$

ou bien d'après les équations (26), les projections x', y', z' de cette vitesse seront:

$$x' = \sum_{\alpha=1}^{k} \left(\frac{\partial x_0}{\partial q_\alpha} + \sum_{i=1}^{p} \frac{\partial x_0}{\partial q_{k+i}} a_{i\alpha} \right) q_\alpha', \qquad y' = \ldots, \qquad z' = \ldots$$

D'où l'on voit que la force vive $2T$ du système

$$2T = \sum m(x'^2 + y'^2 + z'^2),$$

est une fonction quadratique des $q_1', q_2', \ldots, q_k'$

$$2T = \varphi(q_1', q_2', \ldots, q_k').$$

Les projections $(x')_0, (y')_0, (z')_0$ de la vitesse $(v)_0$ avant le choc s'obtiennent en donnant aux q_α' les valeurs $(q_\alpha')_0$; tandis que les projections $(x')_1, (y')_1, (z')_1$ de la vitesse $(v)_1$ après le choc s'obtiennent en donnant aux q_α' les valeurs $(q_\alpha')_1$; alors les projections de la vitesse perdue sont:

$$(x')_0 - (x')_1 = \sum_{\alpha=1}^{k} \left(\frac{\partial x_0}{\partial q_\alpha} + \sum_{i=1}^{p} \frac{\partial x_0}{\partial q_{k+i}} a_{i\alpha} \right) p_\alpha,$$

$$(y')_0 - (y')_1 = \ldots, \quad (z')_0 - (z')_1 = \ldots,$$

en écrivant

$$p_\alpha = (q_\alpha')_0 - (q_\alpha')_1 \qquad (\alpha = 1, 2, \ldots, k). \tag{28}$$

Par conséquent la force vive due aux vitesses perdues

$$2T' = \sum m \{[(x')_0 - (x')_1]^2 + [(y')_0 - (y')_1]^2 + [(z')_0 - (z')_1]^2\}$$

se déduit de l'expression générale $2T$ pour la force vive en y remplaçant $q_1', q_2', \ldots, q_k'$ par $p_1, p_2, \ldots, p_k$.

Calculons maintenant la force vive perdue. D'après le théorème pour les fonctions homogènes, nous aurons

$$(2T)_0 - (2T)_1 = \sum_{\alpha=1}^{k} (q_\alpha')_0 \left(\frac{\partial T}{\partial q_\alpha'} \right)_0 - \sum_{\alpha=1}^{k} (q_\alpha')_1 \left(\frac{\partial T}{\partial q_\alpha'} \right)_1,$$

ce qui peut s'écrire en ajoutant et retranchant la quantité $\sum (q'_\alpha)_1 \left(\frac{\partial T}{\partial q'_\alpha}\right)_0$,

$$(2T) - (2T)_1 = \sum_{\alpha=1}^{k} [(q'_\alpha)_0 - (q'_\alpha)_1] \left(\frac{\partial T}{\partial q'_\alpha}\right)_0 + \sum_{\alpha=1}^{k} (q'_\alpha)_1 \left[\left(\frac{\partial T}{\partial q'_\alpha}\right)_0 - \left(\frac{\partial T}{\partial q'_\alpha}\right)_1\right].$$

En tenant compte de la relation (28), et que

$$\sum_{\alpha=1}^{k} p_\alpha \left(\frac{\partial T}{\partial q'_\alpha}\right)_0 = \sum_{\alpha=1}^{k} (q'_\alpha)_0 \frac{\partial T}{\partial p_\alpha}$$

d'après une propriété connue des formes quadratiques; remarquant enfin que $\frac{\partial T}{\partial p_\alpha}$ est une fonction linéaire et homogène en p_α et que par conséquent

$$\frac{\partial T}{\partial p_\alpha} = \left(\frac{\partial T}{\partial q'_\alpha}\right)_0 - \left(\frac{\partial T}{\partial q'_\alpha}\right)_1,$$

nous obtenons pour la force vive perdue:

$$(29) \qquad (2T)_0 - (2T)_1 = \sum_{\alpha=1}^{k} [(q'_\alpha)_0 + (q'_\alpha)_1] \left[\left(\frac{\partial T}{\partial q'_\alpha}\right)_0 - \left(\frac{\partial T}{\partial q'_\alpha}\right)_1\right]$$

$$= \sum_{f=1}^{n} [(q'_f)_0 + (q'_f)_1] \left[\left(\frac{\partial T}{\partial q'_f}\right)_0 - \left(\frac{\partial T}{\partial q'_f}\right)_1\right]$$

$$+ \sum_{j=1}^{r} [(q'_{n+j})_0 + (q'_{n+j})_1] \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_0 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_1\right].$$

Mais d'après les équations (8') nous avons

$$\left(\frac{\partial T}{\partial q'_f}\right)_0 + \sum_{j=1}^{r} \left(\frac{\partial T}{\partial q'_{n+j}}\right)_0 b_{jf} = \left(\frac{\partial T}{\partial q'_f}\right)_1 + \sum_{j=1}^{r} \left(\frac{\partial T}{\partial q'_{n+j}}\right)_1 b_{jf} \quad (f = 1, 2, \ldots, n)$$

d'où l'on tire

$$(30) \qquad \left(\frac{\partial T}{\partial q'_f}\right)_0 - \left(\frac{\partial T}{\partial q'_f}\right)_1 = \sum_{j=1}^{r} b_{jf} \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_1 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_0\right] \quad (f = 1, 2, \ldots, n).$$

D'autre part les équations (27) nous donnent

$$(31) \qquad (q'_{n+j})_1 = \sum_{f=1}^{n} b_{jf} (q'_f)_1 \qquad (j = 1, 2, \ldots, r).$$

Alors la formule (29), en tenant compte des formules (30) et (31), devient

$$(32) \qquad (2T)_0 - (2T)_1 = \sum_{f=1}^{n} (q'_f)_0 \sum_{j=1}^{r} b_{jf} \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_1 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_0\right]$$

$$+ \sum_{j=1}^{r} (q'_{n+j})_0 \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_0 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_1\right].$$

Calculons maintenant la force vive due aux vitesses perdues. Nous aurons

$$2T' = \sum_{\alpha=1}^{k} p_\alpha \frac{\partial T}{\partial p_\alpha} = \sum_{\alpha=1}^{k} [(q'_\alpha)_0 - (q'_\alpha)_1] \left[\left(\frac{\partial T}{\partial q'_\alpha}\right)_0 - \left(\frac{\partial T}{\partial q'_\alpha}\right)_1\right]$$

$$= \sum_{f=1}^{n} [(q'_f)_0 - (q'_f)_1] \left[\left(\frac{\partial T}{\partial q'_f}\right)_0 - \left(\frac{\partial T}{\partial q'_f}\right)_1\right]$$

$$+ \sum_{j=1}^{r} [(q'_{n+j})_0 - (q'_{n+j})_1] \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_0 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_1\right],$$

ou bien, en tenant compte des équations (30) et (31),

$$(33) \qquad 2T' = \sum_{f=1}^{n} (q'_f)_0 \sum_{j=1}^{r} b_{jf} \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_1 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_0\right]$$

$$+ \sum_{j=1}^{r} (q'_{n+j})_0 \left[\left(\frac{\partial T}{\partial q'_{n+j}}\right)_0 - \left(\frac{\partial T}{\partial q'_{n+j}}\right)_1\right].$$

En comparant (32) et (33) nous obtenons

$$(2T)_0 - (2T)_1 = 2T'.$$

Ce qui est l'expression du théorème de Carnot.

1^er^ décembre 1923.

(Eingegangen am 10. 12. 1923.)

Reduzible Kurven vom Maximalindex.

Von

Hans Mohrmann in Basel.

In seiner jüngsten Annalen-Arbeit [1]) hat Herr Nagy wohl zum ersten Male auch *reduzible* ebene Kurven vom Maximalindex in den Bereich der Betrachtung gezogen, d. h. Kurven, die mit keiner Geraden der projektiven Ebene mehr als n und weniger als $n-2$ reelle Punkte gemein haben. Das einfachste Beispiel einer derartigen Kurve bildet eine Hyperbel mit einer Ellipse, die beide Äste der Hyperbel in reellen Punkten schneidet [2]). Herr Nagy hat sechs Sätze (Nr. XLIV bis XLIX) über diese Kurven formuliert. Die beiden ersten ergeben sich unmittelbar auf Grund der Begriffe der Reduzibilität und Irreduzibilität nicht-algebraischer Kurven und meines Satzes, nach dem eine irreduzible Kurve vom Maximalindex höchstens ein Oval (Kurvenzug 2. Ordnung vom Index 0) besitzen kann. Die beiden folgenden Sätze (XLVI und LXVII) und damit auch die ihnen dualistisch entsprechenden Sätze XI und XIII, aus denen sie gewonnen sind, sind unrichtig [3]). Die beiden letzten Sätze endlich, die sich leicht und in anschaulicher Weise auch direkt beweisen lassen, und zwar mit ausschließlicher Benutzung von Kreisen und den trisymmetrischen Kurven 3. Ordnung mit den Gleichungen in Polarkoordinaten r, φ (vgl. Fig. 1)

$$r = \frac{a}{\sin 3\varphi}, \tag{1}$$

handeln von Kurven mit der Maximalzahl im Endlichen liegender Ovale.

---

[1]) Math. Annalen **89**, S. 32 ff.

[2]) Vgl. Math. Annalen **74**, S 319.

[3]) [Anmerkung vom 10. April 1924.] Sie sind inzwischen von Herrn Nagy (diese Annalen **90**, S. 150—151) berichtigt worden und zwar auf Grund von Mitteilungen, die ihm die Annalen-Redaktion nach Empfang einer kurzen aufklärenden, aber nicht aufgenommenen Bemerkung von mir gemacht hat. Auch das Manuskript der vorliegenden Arbeit hat sich erstmals Anfang Juni 1923 in den Händen der Annalen-Redaktion befunden.

Die Typen solcher Kurven, die Herr Nagy auf Grund des von ihm eingeschlagenen, in den elementarsten Fällen bestechenden dualistischen Verfahrens erhalten hat, erschöpfen nur für die Kurven $(3k-1)$-ter Ordnung $(k=1, 2, 3, \ldots)$ die möglichen Arten. Beim Aufbau der Kurven $3k$-ter und $(3k+1)$-ter Ordnung lassen sich auch irreduzible Kurven 4. und 5. Ordnung verwenden, wodurch die Mannigfaltigkeit ganz außerordentlich wächst. So gibt es z. B. neun wesentlich unterschiedene Arten (reduzibler) Kurven 7. Ordnung (vom Maximalindex) mit zwei im Endlichen liegenden Ovalen. Nur eine von diesen enthält *kein* von Doppelpunkten freies Oval, und diese ist die *einzige*, die Herr Nagy mittels des für seine Existenzbeweise benutzten Verfahrens gewonnen hat.

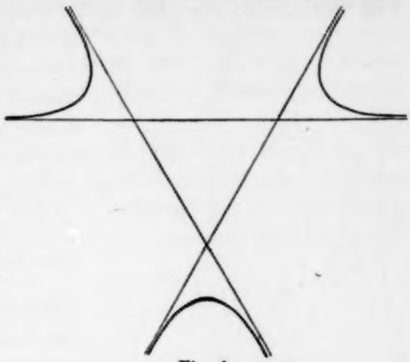

Fig. 1.

Es sei mir daher gestattet, auf die *reduziblen* Kurven vom Maximalindex, deren sämtliche paaren Züge bis auf einen Zug 3. Ordnung Ovale sind, sowie diejenigen mit der Maximalzahl im Endlichen liegender Ovale bei vorgegebener Ordnung kurz erneut einzugehen.

I.

Kurven $2k$-ter und $(2k+1)$-ter Ordnung vom Maximalindex mit k Ovalen. Der Hyperbelprozeß.

Zunächst wollen wir die Kurven vom Maximalindex behandeln, deren sämtliche Züge, bis auf einen (bei Kurven ungerader Ordnung notwendigen) Zug 3. Ordnung Ovale sind. Eines von diesen dürfen wir, vom projektiven Standpunkt aus, immer als im Endlichen liegend voraussetzen. Wir wählen das Oval, von dem wir beim Aufbau der Kurven gerader Ordnung ausgehen, oder das Oval der Ausgangskurve 3. Ordnung bei den Kurven ungerader Ordnung. Das Verfahren, das wir beim Aufbau verwenden, bezeichnen wir als *Hyperbelprozeß*. Es besteht in folgendem:

Ist $\widehat{A_1 A_2}$ ein ganz im Endlichen liegender, überall konvexer Kurvenbogen, dessen Anfangs- und Endtangente einen Winkel von nicht mehr als 180^0 bilden, und H eine algebraische oder nicht-algebraische Hyperbel (Kurvenzug 2. Ordnung vom Index 0 mit zwei getrennten unendlich fernen Punkten), deren einer Ast durch die Punkte A_1 und A_2 hindurchgeht,

während der andere ebenfalls mit dem Bogen $\widehat{A_1 A_2}$ zwei (getrennte) reelle Punkte gemein hat, so bildet die Hyperbel mit dem gegebenen Bogen $\widehat{A_1 A_2}$ nach Fortlassung des ganz im Endlichen liegenden Hyperbelbogens $\widehat{\widehat{A_1 A_2}}$ eine Kurve, die (auf zwei Weisen) in zwei unpaare Züge mit je zwei Ecken zerlegt werden kann. Ist der Kurvenbogen aus einem geschlossenen Kurvenzuge herausgeschnitten, so schließt der Hyperbelbogen $\widehat{\widehat{A_1 A_2}}$ diesen Zug wieder. Gehört der Kurvenbogen $\widehat{A_1 A_2}$ einer Kurve vom Maximalindex an und *ersetzt der Hyperbelbogen* $\widehat{\widehat{A_1 A_2}}$ *den ausscheidenden Kurvenbogen* $\widehat{A_1 A_2}$ *auch hinsichtlich der Maximalindexeigenschaft der übrigbleibenden Kurve,* so entsteht durch den Hyperbelprozeß wieder eine Kurve vom Maximalindex.

Fig. 2.

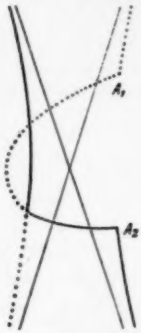

Fig. 3.

Hieraus folgt unmittelbar, daß *aus jeder Kurve n-ter Ordnung vom Maximalindex mit k Ovalen eine (reduzible) Kurve* $(n+2)$*-ter Ordnung mit* $k+1$ *Ovalen gewonnen werden kann, die wieder vom Maximalindex ist.*

Geht man von einem Oval oder von einer Kurve 3. Ordnung mit Oval aus, so erhält man somit durch $(k-1)$-malige passende Anwendung des Hyperbelprozesses Kurven $2k$-ter bzw. $(2k+1)$-ter Ordnung vom Maximalindex, deren sämtliche paaren Züge Ovale sind. Beachtet man noch, daß jede Tangente der Hyperbel den Bogen $\widehat{A_1 A_2}$ in zwei reellen Punkten schneidet, so sieht man, daß die Anzahl der Doppelpunkte der entstehenden Kurve $2k$-ter bzw. $(2k+1)$-ter Ordnung (wofern keine mehrfachen Punkte auftreten) die algebraisch mögliche Maximalzahl erreicht und nicht überschreitet. Sie beträgt demnach[4]):

$$4\binom{k}{2} \text{ bzw. } 4\binom{k-1}{2} + 6(k-1) = 2(k^2-1).$$

Man kann den Hyperbelprozeß, ausgehend von einem Oval oder dem Oval einer Kurve 3. Ordnung, immer wieder auf ein Oval, insbesondere *dasselbe* Oval anwenden. In diesem Falle kann nur ein Oval der Kurve ganz im Endlichen liegen, da einander schneidende Ovale einer Kurve vom Maximalindex niemals beide ins Endliche gebracht werden können. Bei den Kurven $(2k+1)$-ter Ordnung mit k Ovalen kann man aber

---

[4]) Nagy, a. a. O., S. 74.

auch von dem *unpaaren* Zug der Ausgangskurve 3. Ordnung ausgehen und beim Aufbau ausschließlich diesen verwenden. Alsdann entstehen Kurven vom Maximalindex, deren $k-1$ neue Ovale sämtlich das alte gebundene (von Doppelpunkten freie) Oval umschließen. Bei den Kurven *ungerader* Ordnung, deren sämtliche Züge bis auf einen Zug 3. Ordnung Ovale sind, können demnach auch zwei und nur zwei Ovale ganz im Endlichen liegen. So gibt es z. B. zwei verschiedene Typen reduzibler Kurven 5. Ordnung mit zwei Ovalen.

Indessen brauchen keineswegs *alle* paaren Züge einer Kurve $2k$-ter oder $2k+1$-ter Ordnung mit k paaren Zügen Ovale zu sein. Ein einfaches „Gegenbeispiel" liefert die Kurve 6. Ordnung, die aus dem reellen Zuge der unicursalen Kurve 4. Ordnung mit der Gleichung in Polarkoordinaten r, φ (vgl. Fig. 4)

$$r = \frac{2}{\sin 2\varphi}, \tag{2}$$

dem Kreise mit dem Mittelpunkte jener Kurve (2) als Mittelpunkt und dem Radius r_1 $(0 < r_1 < 1)$ sowie dem konzentrischen Kreise mit dem Radius $r_2 > 2$ besteht[5]). Die Kurve (2) hat in rechtwinkligen kartesischen Koordinaten $x = r\cos\varphi$ und $y = r\sin\varphi$ die Gleichung

$$\frac{1}{x^2} + \frac{1}{y^2} = 1 \tag{2*}$$

und ist unter dem Namen „Kreuzkurve" wohlbekannt. Sie hat vier Symmetrieachsen, die sich in einem Punkte, dem Koordinatenanfang, schneiden, der für die Kurve isolierter Doppelpunkt (und Mittelpunkt) ist. Der reelle Zug der Kurve ist 4. Ordnung und vom Index 2. Er hat zwei Doppelpunkte, die im Unendlichen liegen und doppelte Wendepunkte für die Kurve sind. Die vier Wendeasymptoten bilden ein Quadrat mit der Seitenlänge 2. Ersetzt man den isolierten Doppelpunkt M der Kurve durch ein Oval, speziell einen Kreis k_1 vom Radius r_1 $(0 < r_1 < 1)$, mit M als Mittelpunkt, so erhält man eine *irreduzible* (nicht-algebraische) Kurve 4. Ordnung vom Maximalindex mit Oval, die wir weiterhin als C_4 bezeichnen wollen. Fügt man der C_4 noch den Kreis k_2 mit M als Mittelpunkt und $r_2 > 2$ als Radius hinzu, so entsteht eine *reduzible* Kurve 6. Ordnung C_6 *vom Maximalindex*. Denn da der Kreis k_2 den paaren Zug 4. Ordnung der C_4 in acht (reellen) Punkten schneidet, so kann man diesen Zug zusammen mit k_2 so zerschneiden, daß vier *unpaare* Züge (mit je zwei Ecken) entstehen (vgl. Fig. 5). Die C_6 besitzt drei paare Züge, von

[5]) Beliebig viele weitere Beispiele ergeben sich aus den Prozessen des (nächsten) II. Abschnitts. — Alle Figuren in dieser Arbeit sind *schematisch*!

denen nur zwei Ovale sind. Wir vermerken noch, daß auch die Zahl der Doppelpunkte der Kurven $2k$-ter oder $2k+1$-ter Ordnung mit k paaren

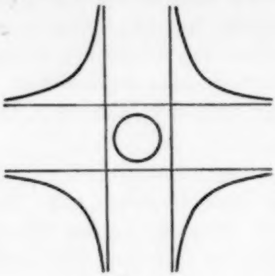

Fig. 4.

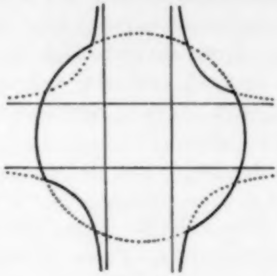

Fig. 5.

Zügen eine Modifikation erleidet, wenn nicht alle paaren Züge Ovale sind Die soeben erzeugte C_6 z. B. hat nur zehn Doppelpunkte gegenüber zwölf der aus drei Ovalen bestehenden C_6 vom Maximalindex.

II.

Die Kurven gegebener Ordnung vom Maximalindex mit der Maximalzahl im Endlichen liegender Ovale.

Da eine irreduzible Kurve vom Maximalindex nur ein Oval besitzen kann und andrerseits der Hyperbelprozeß nur Kurven vom Maximalindex mit *höchstens zwei* im Endlichen liegenden Ovalen erzeugt, so können am Aufbau reduzibler Kurven mit k im Endlichen liegenden Ovalen höchstens zwei „Kegelschnitte" beteiligt sein, wenn k den größtmöglichen Wert bei vorgegebener Ordnung annehmen soll. Ist dies der Fall, so ist die Ordnung notwendig (mindestens) $3k+1$. Ist nur ein irreduzibler Bestandteil der Kurve mit k im Endlichen liegenden Ovalen ein Kegelschnitt, so kann die Ordnung auch $3k$ und $3k-1$ betragen. In letzterem Falle ist *notwendig* ein irreduzibler Bestandteil ein Kegelschnitt.

Daß wirklich Kurven vom Maximalindex der Ordnungen $3k \pm \varepsilon$ $(\varepsilon = 0, 1)$ mit k im Endlichen liegenden Ovalen existieren, hat Herr Nagy, von den Kurven vom Maximalklassenindex ausgehend, erwiesen. Ich führe hier den Beweis direkt und gebe zugleich eine *erschöpfende Klassifikation der möglichen* Fälle.

Die Gleichung in Polarkoordinaten r, φ

$$(1^*) \qquad r = \frac{3}{\sin 3\varphi}$$

stellt eine Kurve 3. Ordnung mit drei Symmetrieachsen dar, die sich im Nullpunkt schneiden, der für die Kurve isolierter Doppelpunkt und Mittel-

punkt ist. Die drei (reellen) Wendepunkte der Kurve liegen im Unendlichen, und die drei Wendeasymptoten bilden ein reguläres Dreieck von der Höhe 3, dessen Schwerpunkt im Mittelpunkt M der Kurve liegt. Ersetzt man den isolierten Doppelpunkt M durch einen Kreis k_1 mit M als Mittelpunkt und r_1 $(0 < r_1 < 1)$ als Radius, so entsteht eine *irreduzible* (nicht-algebraische) Kurve 3. Ordnung mit Oval. Wir nennen sie C_3 (vgl. Fig. 6).

Fügt man der C_3 einen Kreis k_3 mit M als Mittelpunkt und $r_3 > 3$ als Radius hinzu, so erhält man eine *reduzible* Kurve 5. Ordnung C_5 vom Maximalindex (mit zwei im Endlichen liegenden Ovalen). Denn man kann den unpaaren Zug der C_5 zusammen mit k_3 (auf verschiedene Weisen) so zerschneiden, daß drei unpaare Züge mit je zwei Ecken entstehen, wodurch der Maximalindex verbürgt ist.

Fügt man der C_3 eine zu ihr ähnliche C_3' hinzu, für die der Kreis k_1' mit dem soeben genannten Kreise k_3 identisch ist, so erhält man eine reduzible Kurve 6. Ordnung vom Maximalindex mit zwei im Endlichen liegenden Ovalen: Fig. 6.

Fügt man der C_3 eine der beschriebenen C_4 ähnliche C_4' hinzu, für die der Kreis k_1' mit dem genannten Kreise k_3 identisch ist, so erhält man eine reduzible (nicht-algebraische) Kurve 7. Ordnung vom Maximalindex mit zwei im Endlichen liegenden Ovalen.

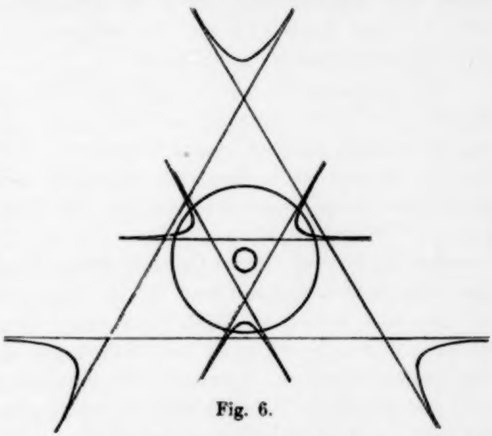

Fig. 6.

Entsprechendes gilt für die analoge Kombination einer C_4 mit einer C_3'. Während aber bei der ersten Kurve 7. Ordnung auf dem einen Oval der Kurve sechs Doppelpunkte liegen, enthält das eine Oval der zweiten Kurve 7. Ordnung acht Doppelpunkte. Man gewinnt so zwei topologisch unterschiedene Typen von Kurven 7. Ordnung, vom Maximalindex mit zwei im Endlichen liegenden Ovalen, denen man übrigens eine Reihe weiterer hinzugesellen kann (siehe III. Abschnitt).

Durch analoge Kombination einer C_4 mit einer C_4' entsteht eine reduzible Kurve 8. Ordnung vom Maximalindex, die vier *paare Züge* aufweist, von denen jedoch *nur* zwei *Ovale* sind usw.

Beliebige Kombinationen der genannten Prozesse ergeben immer Kurven vom Maximalindex mit lauter im Endlichen gelegenen Ovalen, die hier sämtlich konzentrische Kreise sind. An diesen Prozessen können beliebig viele C_3 und C_4, aber jeweils nur höchstens *ein* k_3, d. h. ein nicht gebundenes Oval C_2 beteiligt sein. In jedem Falle erhält man Kurven mit r Zügen, von denen, je nachdem r gerade oder ungerade ist, $\frac{r}{2}$ oder $\frac{r+1}{2}$ im Endlichen gelegene Ovale sind. Und es ist auf Grund des über den Hyperbelprozeß Gesagten klar, daß auch nicht mehr Ovale einer r-zügigen Kurve vom Maximalindex im Endlichen liegen können[6]).

Eine derartige Kurve, an deren Bildung αC_2 $(\alpha = 0, 1)$, βC_3, γC_4 $(\beta, \gamma = 1, 2, 3, \ldots)$ der beschriebenen Art beteiligt sind, ist von der Ordnung

$$2\alpha + 3\beta + 4\gamma$$

und enthält $\alpha + \beta + \gamma$ im Endlichen liegende Ovale.

Hieraus folgen die Sätze: *Die niedrigste Ordnung einer (reduziblen) Kurve vom Maximalindex mit k im Endlichen liegenden Ovalen beträgt* $3k-1$. *Eine Kurve $(3k \pm \varepsilon)$-ter Ordnung $(\varepsilon = 0, 1)$ kann höchstens k im Endlichen liegende Ovale haben.*

Die Maximalzahl der im Endlichen liegenden Ovale kann bei einer Kurve $(3k-1)$-ter Ordnung im wesentlichen nur auf *eine* Weise erreicht werden, nämlich dadurch, daß $k-1$ Kurven 3. Ordnung C_3 und ein freies Oval k_3 benutzt wird. Letzteres umschließt notwendig sämtliche $k-1$ gebundenen Ovale, deren innerstes frei von Doppelpunkten ist. Es trägt $6(k-1)$ Doppelpunkte.

Bei den Kurven $3k$-ter Ordnung können lauter C_3 oder auch *eine* C_2 *und* dann notwendig auch eine C_4 der beschriebenen Art *oder* auch eine C_4^* mit *zwei unpaaren* Zügen (und einem Oval) benutzt werden. Das gibt $1 + 2(k-1) = 2k-1$ *hinsichtlich den Kurvenzügen oder den auf den Ovalen liegenden Doppelpunkten verschiedene Typen. Es gibt also z. B. drei wesentlich unterschiedene Kurven* 6. *Ordnung vom Maximalindex mit zwei im Endlichen liegenden Ovalen.*

An der Bildung der Kurven $(3k+1)$-ter Ordnung endlich können außer C_3 eine oder zwei C_4 bzw. C_4^* beteiligt sein. In den Fällen, in denen zwei Kurven 4. Ordnung irreduzible Bestandteile sind, ist notwendig auch eine irreduzible Kurve 2. Ordnung vorhanden. Man gewinnt auf diese Weise

$$2k + 2\frac{(k-1)!}{2!(k-3)!} + \frac{(k-1)!}{(k-3)!} = 2k + 4\binom{k-1}{2}$$

wesentlich unterschiedene Arten.

---

[6]) Nagy, a. a. O. Satz XLIX.

Man kann aber zur Erzeugung von Kurven $(3k+1)$-ter Ordnung auch irreduzible Kurven 5. Ordnung verwenden, die vom Maximalindex sind und ein Oval besitzen. Solcher Kurven gibt es drei verschiedene Arten: Kurven mit drei unpaaren Zügen und einem Oval; Kurven mit einem unpaaren Zug, einem paaren Zug 4. Ordnung und einem Oval; und endlich Kurven mit einem unpaaren Zug 5. Ordnung und einem Oval. Die ersten beiden Arten irreduzibler Kurven 5. Ordnung vom Maximalindex gewinnt man leicht aus einer C_3 oder einer C_4, indem man ihr zwei oder einen passend gewählten Kurvenzug 3. Ordnung zusetzt, z. B. unter Benutzung des reellen Zuges der Kurve 3. Ordnung

$$(3) \qquad y = \frac{c}{1+x^2},$$

der mit hinreichend kleinem c einer Geraden sehr nahe kommt. Beispielsweise erhält man aus der früher beschriebenen C_3 eine C_4 mit zwei unpaaren Zügen und einem Oval, wenn man der C_3 mit dem reellen Zuge der Kurve

$$(1^*) \qquad r = \frac{3}{\sin 3\varphi}$$

den reellen Zug der Kurve

$$(3^*) \qquad y+1 = \frac{-1}{1+x^2}$$

zufügt (Fig. 7). Setzt man noch den Kurvenzug dritter Ordnung hinzu, der aus (3*) durch Spiegelung an der Geraden

$$y = \frac{1}{3}\sqrt{3}\,x$$

hervorgeht, so gewinnt man eine irreduzible Kurve 5. Ordnung vom Maximalindex mit drei unpaaren Zügen und einem Oval, die sich ebenfalls unserem Aufbauprozeß der Kurven mit lauter im Endlichen liegenden Ovalen anschaulich eingliedert.

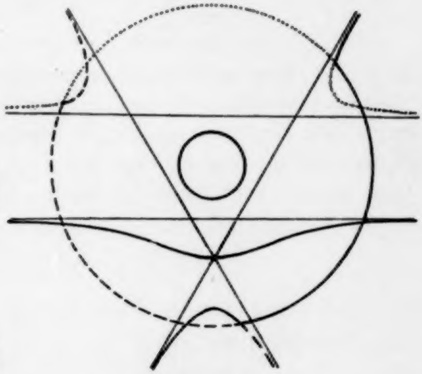

Fig. 7.

Ganz analog gewinnt man aus den beschriebenen C_4 mit einem *paaren* Zuge 4. Ordnung durch Zusatz eines unpaaren Zuges 3. Ordnung in passender Lage eine irreduzible Kurve 5. Ordnung vom Maximalindex mit einem unpaaren Zuge 3., einem paaren Zuge 4. Ordnung und einem Oval, die sich ebenfalls unserem Erzeugungsprozeß ohne weiteres eingliedert.

Das gleiche gilt von der irreduziblen Kurve 5. Ordnung mit einem unpaaren Zuge 5. Ordnung vom Maximalindex und einem Oval, wie man

sie genau wie früher unter Verwendung des reellen Zuges der unikursalen Kurve 5. Ordnung mit der Gleichung in Polarkoordinaten r, φ

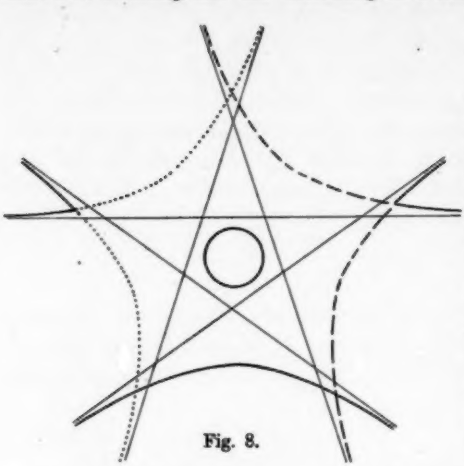

Fig. 8.

(4) $$r = \frac{\frac{5}{3}}{\sin\frac{5}{3}\varphi}$$

erhält, wenn man diesem einen Kreis k_1 mit r_1 als Radius ($0 < r_1 < 1$) und dem isolierten Doppelpunkt der Kurve (4) als Mittelpunkt hinzufügt (Fig. 8). Durch Hinzunahme eines zweiten, konzentrischen Kreises k_4 mit dem Radius $r_4 > \frac{10}{3}$, der also die fünf Selbstdurchsetzungspunkte der Kurve 5. Ordnung einschließt, gewinnt man eine reduzible Kurve 7. Ordnung, die wiederum vom Maximalindex ist, da man den Zug 5. Ordnung zusammen mit k_4 leicht (auf verschiedene Weisen) in fünf unpaare Züge mit Ecken zerschneiden kann. Man braucht dabei nur zu beachten, daß sich der Zug 5. Ordnung in drei unpaare Züge zerlegen läßt (vgl. Fig. 8), von denen zwei je eine Ecke und je einen unendlich fernen Punkt besitzen, während der dritte zwei Ecken aufweist und vom Typus des reellen Zuges der Kurve

(1) $$r = \frac{a}{\sin 3\varphi}$$

ist. Auch diese irreduzible Kurve 5. Ordnung vom Maximalindex mit einem Zuge 5. Ordnung und einem Oval gliedert sich also unserem Erzeugungsprozeß völlig ein, so daß zu den schon angegebenen Arten, da jeweils nur eine der drei Typen irreduzibler C_5 beteiligt sein kann, noch

$$3(k-1)$$

hinzutreten.

Allen diesen $2k^2 - k + 1$ Arten von reduziblen Kurven $(3k+1)$-ter Ordnung (vom Maximalindex mit k im Endlichen liegenden Ovalen) gemeinsam ist, daß sie *kein weiteres* Oval besitzen, und daß bei jeder *ein Oval frei von Doppelpunkten* ist. Die möglichen Arten sind damit indes noch immer nicht erschöpft. Denn man gewinnt natürlich auch durch einmalige passende Anwendung des *Hyperbelprozesses* aus den Kurven $(3k-1)$-ter Ordnung vom Maximalindex unseren früheren Ausführungen zufolge Kurven $(3k+1)$-ter Ordnung vom Maximalindex mit k im End-

lichen liegenden Ovalen. Diese Kurven besitzen noch ein $(k+1)$-tes Oval, das indessen nicht auch noch ins Endliche gebracht werden kann. Man erkennt leicht, daß es k hinsichtlich den auf den k Ovalen liegenden Doppelpunkten verschiedene Arten dieser Kurven gibt, die man aus den Kurven $(3k-1)$-ter Ordnung dadurch erhält, daß man den Hyperbelprozeß passend[7]) auf eines der k Ovale anwendet. *Nur wenn man den Hyperbelprozeß auf das innerste Oval anwendet, erhält man eine Kurve, die kein von Doppelpunkten freies Oval besitzt. Diese Art ist die einzige der $2k^2+1$ hinsichtlich den Zügen und den auf den Ovalen gelegenen Doppelpunkten verschiedenen Arten von Kurven $(3k+1)$-ter Ordnung vom Maximalindex mit k im Endlichen liegenden Ovalen, die in den Existenzbeweisen von Herrn Nagy auftritt.*

Irreduzible Kurven 6. oder höherer Ordnung können *nicht* Bestandteile einer Kurve vom Maximalindex gegebener Ordnung mit der Maximalzahl im Endlichen liegender Ovale sein. Die vorstehende Klassifikation ist somit *erschöpfend*.

III.

Die reduziblen Kurven 7. Ordnung vom Maximalindex mit (der Maximalzahl) zwei im Endlichen liegenden Ovalen.

Um einen guten Einblick in die Mannigfaltigkeit der Arten von Kurven $(3k+1)$-ter Ordnung vom Maximalindex mit der Maximalzahl k im Endlichen liegender Ovale zu gewinnen, wählen wir die Kurven 7. Ordnung, bei denen nur die Kombinationen mit *zwei* irreduziblen Kurven 4. Ordnung noch nicht möglich sind, die aber alles übrige Wesentliche schon gut erkennen lassen.

Es gibt hier

$$[2k^2+1]_{k=2}=9$$

verschiedene Typen, über die die folgende *Tabelle* die beste Übersicht gibt. In dieser bedeuten:

n_i die *Ordnung* der irreduziblen Kurve, an die das *innere* Oval gebunden ist,

n_a diejenige der Kurve des *äußeren* Ovals;

g die Anzahl der *paaren*,

u diejenige der *unpaaren Züge* der Kurve;

---

[7]) S. Abschnitt I dieser Note. — Numeriert man die Ovale und entsprechend die Kurven 3. Ordnung, an die sie gebunden sind, indem man das m Ovale umschließende Oval als $(m+1)$-tes bezeichnet, so kann man sagen, daß es auf dasselbe hinaus kommt, ob man den Hyperbelprozeß passend auf das $(m+1)$-te Oval oder den unpaaren Zug der m-ten Kurve 3. Ordnung anwendet.

i die Anzahl der *Doppelpunkte* auf dem *inneren Oval*,
a diejenige der *Doppelpunkte* auf dem *äußeren Oval*,
d die *Gesamtzahl der Doppelpunkte* (wenn keine mehrfachen Punkte auftreten), und
w die Anzahl der *Wendepunkte* der Kurve:

Nr.	n_i	n_a	i	a	g	u	d	w
1	4	3	0	8	3	1	14	7
2	3	4	0	6	3	1	14	7
3	4	3	0	8	2	3	13	9
4	3	4	0	6	2	3	13	9
5	5	2	0	10	2	3	13	9
6	5	2	0	10	3	1	14	7
7	5	2	0	10	2	1	15	5
8	3	2	0	10	3	1	16	3
9	3	2	4	10	3	1	16	3

Nur die beiden letzten Arten sind doppelt reduzibel; sie zerfallen in je drei irreduzible Kurven, eine C_3 mit Oval und zwei Kegelschnitte, von denen nur der eine mit dem Oval gleichzeitig im Endlichen liegen kann. Im Falle 8 sind die beiden Kegelschnitte in keiner Weise vor einander ausgezeichnet. Im Falle 9 hingegen schneidet der eine Kegelschnitt das (an die Kurve 3. Ordnung) gebundene Oval. Dieser Kegelschnitt kann weder mit dem Oval noch mit dem anderen Kegelschnitt gleichzeitig ins Endliche gebracht werden. Die Anzahl der Doppelpunkte der Gesamtkurve, die auf dem Kegelschnitt liegen, ist, wie bei der andern, 10; doch gehören hier nur zwei dieser Doppelpunkte dem unpaaren Zuge der Kurve 3. Ordnung an, während der andere ihn in sechs Punkten schneidet.

Art 9 ist die einzige Kurve, die kein von Doppelpunkten freies Oval besitzt.

Basel, den 2. Juni 1923.

(Eingegangen am 1. 10. 1923.)

Über die Abbildung der projektiven Ebene auf eine geschlossene singularitätenfreie Fläche im erreichbaren Gebiet des Raumes.

Von

Friedrich Schilling in Danzig-Langfuhr.

Auf Anregung von Herrn D. Hilbert hat W. Boy zum ersten Male die interessante Aufgabe gelöst, die projektive Ebene punktweise eineindeutig abzubilden auf eine geschlossene, singularitätenfreie Fläche, die in einem erreichbaren Gebiet des Raumes gelegen ist (z. B. innerhalb einer Kugel mit gegebenem Radius um einen uns vor Augen liegenden Punkt des Raumes)[1]. Herr Boy hat zur Lösung der Aufgabe sehr viel weiter reichende Untersuchungen der Analysis situs benutzt; auch läßt seine Methode es vermissen, daß die Abbildung der projektiven Ebene selbst auf die Fläche hinsichtlich der Zuordnung der einzelnen Punkte erfaßt wird.

Wir wollen daher im folgenden diese Abbildung nach einer neuen, überraschend einfachen Methode durchführen, die so elementar und an-

---

[1]) Werner Boy, *Über die Curvatura integra und die Topologie geschlossener Flächen*, Dissertation, Göttingen 1901, abgedruckt in den Math. Annalen 57, S. 151 ff. Im Auszug gibt das Resultat seiner Untersuchungen die Arbeit: *Über die Abbildung der projektiven Ebene auf eine im Endlichen geschlossene, singularitätenfreie Fläche*, Göttinger Nachrichten 1901, S. 1–14. Die von Boy konstruierten zugehörigen beiden Modelle sind gleichzeitig im Verlage von Martin Schilling in Leipzig als Serie XXX Nr. 1 und 2 erschienen.

Im Gegensatz zu Boy habe ich im Text statt „im Endlichen“ geschlossene Fläche ausdrücklich „Fläche in einem erreichbaren Gebiet des Raumes“ gesagt, um damit sogleich anzudeuten, daß nicht nur für Euklidische, sondern auch für elliptische und hyperbolische Geometrie unsere Aufgabe formuliert sein soll. In der elliptischen Geometrie würde ja die projektive Ebene selbst schon die Forderung erfüllen, eine im Endlichen geschlossene, singularitätenfreie Fläche zu sein.

Wegen der weiteren Literatur sei auf die Arbeit von Boy verwiesen. Die von uns konstruierte Fläche wird man leicht mit der ersten Fläche von Boy vergleichen können.

schaulich ist, daß sie ohne weiteres in jedem Lehrbuch der projektiven Geometrie Aufnahme finden kann. Unsere Methode benutzt die einfache Regel: Divide et impera, d. h. wir zerschneiden die projektive Ebene $\mathfrak{E}$ in einzelne Teile, verbiegen und verzerren diese in geeigneter Weise und setzen dann diese Teile wieder zu einer geschlossenen Fläche $\mathfrak{F}$ zusammen. Wir wollen unserer Betrachtung wegen der Einfachheit der Darstellung die Verhältnisse der Euklidischen Geometrie zugrunde legen, so daß wir als „uneigentliche Punkte" nur die Punkte der unendlich fernen Geraden zu den eigentlichen Punkten zu adjungieren haben. Doch wird jeder, der mit der modernen nicht-Euklidischen Geometrie vertraut ist, leicht erkennen, daß unsere Ausführungen, sinngemäß übertragen, ebenso für die elliptische wie die hyperbolische Geometrie gelten oder auch für die rein projektive Geometrie, deren Aufbau überhaupt noch kein Parallelenaxiom zugrunde liegt.

Wir projizieren zunächst die horizontal gedachte projektive Ebene $\mathfrak{E}$ auf die von dem horizontalen größten Kreise begrenzte untere Hälfte einer die Ebene berührenden Kugel von deren Mittelpunkt M aus (Fig. 1). Jeder im Endlichen gelegene Punkt P_0 der Ebene findet dann sein Abbild in *einem* bestimmten Punkt P_1 der Halbkugel. Nur jeder unendlich

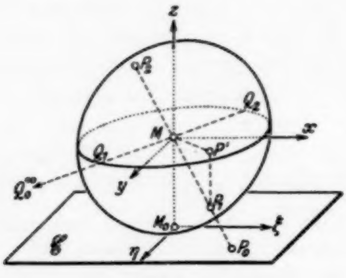

Fig. 1.

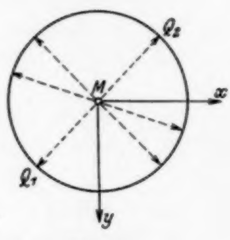

Fig. 2.

ferne Punkt Q_0^∞ der Ebene $\mathfrak{E}$ hat als Abbild zwei diametrale Punkte Q_1, Q_2 auf dem Rande der Halbkugel. Diese Halbkugel sei dann weiter auf die von ihrem Rande begrenzte Kreisfläche orthogonal projiziert. Diese Kreisfläche stellt dann in analoger Weise wie die Halbkugel die Abbildung der ganzen Ebene $\mathfrak{E}$ vor (Fig. 2). Denken wir noch den Rand der Kreisfläche sich selbst der Art zugeordnet, daß je zwei diametrale Punkte einander entsprechen, so ist die Kreisfläche dadurch (analog wie bei den Fundamentalbereichen der automorphen Funktionen) zu einer geschlossenen Fläche geworden.

Es stellt also eine solche Kreisfläche mit der Zuordnung des Randes gewiß das einfachste erreichbare, singularitätenfreie Abbild der projektiven Ebene dar. Dieses Abbild ist auch stets ebenso zu verwenden, wie die später von uns konstruierte geschlossene Fläche $\mathfrak{F}$[2]).

Folgende kleine etwas abschweifende Bemerkung sei noch einzufügen gestattet: Wir denken diese Kreisfläche so auf die Ebene $\mathfrak{E}$ gelegt, daß ihr Mittelpunkt M mit dem Berührungspunkt M_0 der Ebene $\mathfrak{E}$ und der Kugel zusammenfällt und die Durchmesser der Kreisfläche mit den entsprechenden Strahlen durch M_0. Wird dann diese so entstandene Figur in der Ebene $\mathfrak{E}$ um irgendeinen Strahl dieser Ebene durch den Punkt M_0 gedreht, so erhalten wir aus der Kreisfläche eine Kugel, deren diametrale Punkte einander zugeordnet sind, und der von dieser Kugel eingeschlossene Raum ist das Abbild des ganzen durch die Drehung aus der Ebene $\mathfrak{E}$ entstandenen projektiven Raumes. Somit kann man also die Euklidische, elliptische und hyperbolische Geometrie des Raumes mit allen eigentlichen und uneigentlichen Punkten, Geraden und Ebenen, mit der absoluten Fläche 2. Grades und dem zugehörigen absoluten Polarsystem in diesem durch die Zuordnung der Begrenzung geschlossenen Kugelraum deuten, insbesondere hinsichtlich eines anschaulichen axiomatischen Aufbaus dieser Geometrien, worauf wir jedoch hier nicht näher eingehen können.

Jetzt soll nun aber aus der Kreisfläche der Fig. 2 eine neue Fläche $\mathfrak{F}$, natürlich unter Verzerrungen, zusammengebogen werden, indem man den Rand der Zuordnung entsprechend mit sich selbst vereinigt. Während wir also bisher projektiv verfahren sind, kommen nun die Anschauungen der Analysis situs zu ihrem Recht, in der wir uns Kurven und Flächen gleichsam aus beliebig zusammen- oder ausziehbarem Gummi zu denken haben. Die Verhältnisse sind vergleichsweise denen ganz analog, die beim Zusammenbiegen der Fläche eines Rechtecks mit zugeordneten Rändern gemäß der Fig. 3 zu einer Kreisringfläche vorliegen.

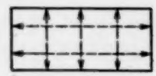

Fig. 3.

Zur Erreichung unseres Zieles zerlegen wir zunächst die Kreisfläche der Fig. 2 durch zwei vom Mittelpunkt M gleich weit abstehende Parallele, deren Abstand wir später noch festlegen werden, in drei Teile (Fig. 4). Diesen beiden Parallelen entsprechen übrigens in der Ebene $\mathfrak{E}$ ersichtlich

[2]) Man kann natürlich die untere Halbkugel auf die Ebene des begrenzenden Kreises auch von dem höchsten Punkt der Kugel aus *stereographisch* projizieren. Dann erhält man als Abbild der projektiven Ebene wieder eine Kreisfläche mit diametral zugeordnetem Rande. Die Geraden der projektiven Ebene sind dann durch Kreisbogen abgebildet, welche die Begrenzung der Kreisfläche in diametralen Punkten schneiden.

die beiden Äste einer Hyperbel (vgl. Fig. 7 und Fig. 12). Der mittlere schraffierte Teil I der Fig. 4 läßt sich nun durch Vereinigen der Randstücke $A_1 C_1$ und $A_2 C_2$ in ein Möbiussches Blatt, wie es durch die sogleich noch näher erläuterte Fig. 5 dargestellt sei, zusammenbiegen, das

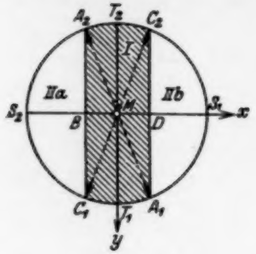

Fig. 4.

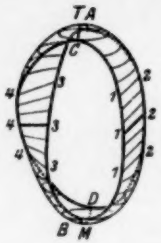

Fig. 5.

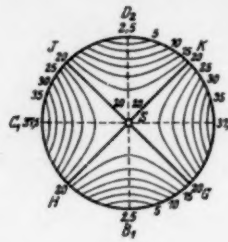

Fig. 6.

ja eine „einseitige“ Fläche mit *einer* Randkurve ist[3]). Es ist hiermit zugleich in einfachster Weise gezeigt:

Die projektive Ebene stellt eine einseitige Fläche vor.

Die beiden anderen Teile IIa und IIb der Fig. 4 können wir durch Vereinigen der beiden begrenzenden Kreisbogen unter entsprechenden Verzerrungen zu einem Vollkreis II vereinigen (Fig. 6). Dies sei wie folgt ausgeführt: Wir denken uns die entsprechenden Gebiete IIa und IIb auf der Halbkugel (Fig. 7). Dann können wir ja das Gebiet IIb durch das diametrale Gebiet II'b ersetzen, das mit IIa zusammen einen Kugelabschnitt bildet. Dieser geht dann durch orthogonale Projektion auf die Ebene seines begrenzenden Kreises in einen Vollkreis über.

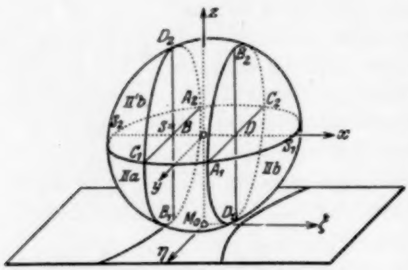

Fig. 7.

Unsere Aufgabe besteht nun darin, diesen Vollkreis und das Möbiussche Blatt mit entsprechenden Randstücken aneinander zu heften. Wir denken zu dem Zweck das Möbiussche Blatt (Fig. 5) der besseren Übersicht wegen für einen Augenblick noch in zwei Teile zerlegt, indem wir

[3]) Vgl. hinsichtlich der einseitigen Flächen z. B. Boy, Dissertation S. 34 ff. oder Math. Annalen l. c. S. 167 ff. oder Möbius, Gesammelte Werke II, Leipzig 1886, S. 484 und S. 519.

es von A über T nach C und von B über M nach D durch Linien zerschneiden, die den Rändern $A_1 C_1$ und $A_2 C_2$ bzw. der Strecke BD in Fig. 4 entsprechen sollen (Fig. 8a, b). Und zwar möge man sich die Entstehung der Teile dann folgendermaßen veranschaulichen: Man denke sich eine Ellipse mit vertikaler großer Achse MT. (Eine Ellipse statt eines Kreises ist nur deswegen gewählt, damit die Figuren, insbesondere Fig. 9, nicht unnötig groß werden.) Längs der einen durch die Punkte M, T begrenzten Hälfte der Ellipse möge eine Strecke so entlang gleiten, daß ihr Mittelpunkt auf der Ellipse bleibt und die Strecke selbst stets zur Ellipsenebene senkrecht steht. Die Strecke beschreibt dann also ein Stück eines elliptischen Zylinderringes (Fig. 8a). Längs der anderen Ellipsenhälfte soll dagegen eine gleichgroße Strecke, auch stets horizontal bleibend, mit ihrem Mittelpunkt auf der Ellipse von derselben Anfangslage durch den Punkt M aus entlang gleiten, doch jetzt so, daß die Strecke sich proportional dem Höhersteigen dreht und zwar insgesamt um 180°. Dieser zweite Teil der Fläche ist also ein Stück einer Art Schraubenfläche (Fig. 8b), deren Gleichung leicht aufzustellen ist. Sind nun diese beiden Flächenteile zusammengesetzt, so daß ihre Ellipsenhälften sich ergänzen, dann sollen noch je der obere und der untere Teil dieser Fläche so durch Zusammendrücken verbogen werden, daß der höchste und der tiefste Punkt elliptische Flächenkrümmung bekommen und auch die beiden Teile ohne Knick sich aneinandersetzen (Fig. 5). Es sei gleich hier noch bemerkt: Den Teil der Fig. 8b kann man als Stück von einer rechts oder links gewundenen „Schraubenfläche" annehmen. Demgemäß erhält man zwei verschiedene Möbiussche Blätter und daraus auch zwei verschiedene geschlossene Flächen $\mathfrak{F}$ und $\mathfrak{F}'$ für die projektive Ebene $\mathfrak{E}$. Diese verschiedenen Flächen sind jedoch als spiegelbildlich zueinander anzusehen, wie ein Paar Handschuhe, und es kann das eine Gebilde durch Umkrempelung in das andere übergeführt werden, indem man zunächst ein kleines Loch in die eine Fläche $\mathfrak{F}$ eingeschnitten denkt und diese Fläche durch das Loch hindurchstülpt und dann das Loch wieder richtig schließt. Wir bevorzugen im folgenden eine *rechts*gewundene Fläche, wie sie auch die Fig. 8b darstellt.

Fig. 8b.

Fig. 8a.

Um nun die Verhältnisse bei der immerhin komplizierten Fläche $\mathfrak{F}$ leicht anschaulich uns vorstellen zu können, wollen wir weiterhin die „kotierte Projektion der Fläche" benutzen. Wie ein Gelände durch seine

Höhenschnitte (Isohypsen) in einer Karte dargestellt ist, so wollen wir auch hier das Möbiussche Blatt und die sich daraus ergebende Fläche $\mathfrak{F}$ durch in gleichem Abstande aufeinander folgende Horizontalschnitte dargestellt denken. Wir können etwa annehmen, daß diese Schnitte im Abstande von 2,5 mm aufeinander folgen. Dann würden als „Koten“ oder Höhenziffern $\varkappa$ zu den gewählten 17 Schnitten hinzuzuschreiben sein: 0; 2,5; 5,0; ...; 40 mm, wobei wir die Höhe h der ganzen Fläche eben zu 40 mm angenommen haben (Fig. 9, entsprechend den verkleinert wiedergegebenen Fig. 5 und 8a, b).

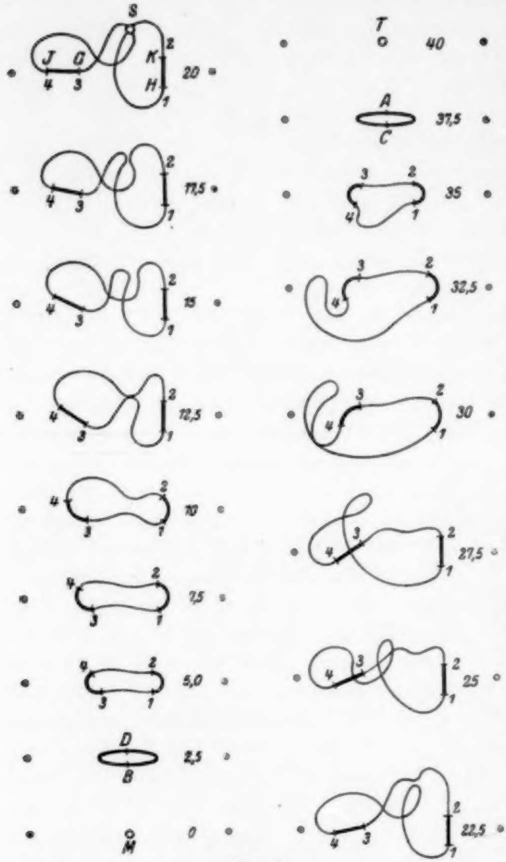

Fig. 9.

Der besseren Übersicht wegen sind die einzelnen Horizontalschnitte übereinander gezeichnet; doch sind jedesmal an beiden Seiten durch kleine Kreise zwei Punkte angegeben, die beim richtigen Übereinanderlegen der Figuren im Raume in zwei senkrechten Geraden zu liegen kommen. In diesen Figuren geben die stark ausgezogenen Bogen oder Strecken 1—2 und 3—4 für sich die Höhenlinien für die beiden Teile der Fig. 5 von dem Möbiusschen Blatt wieder und die schwach ausgezogenen Bogen die Höhenlinien für die verzerrte Kreisfläche der Fig. 6. Diese letzteren Höhenlinien sollen nun den in Fig. 6 eingezeichneten Kurven entsprechen, die wir speziell als Hyperbeläste wählen, worauf wir sogleich zurückkommen.

Beim Übergang von dem Schnitt 12,5 mm zu dem Schnitt 15 mm sind zwei Doppelkurven der Fläche $\mathfrak{F}$ entstanden, was keine Singularität bedingt. Analoges gilt beim Übergang vom Schnitt 22,5 mm zum Schnitt 25 mm. Die vier Schnittpunkte mit den Doppelkurven verschwinden dann wieder beim Übergang vom Schnitt 25 mm zum Schnitt 27,5 mm und von diesem zum Schnitt 30 mm. Besonders zu beachten ist jedoch der Übergang vom Schnitt 17,5 mm über den Schnitt 20 mm zum Schnitt 22,5 mm. Beim Schnitt 20 mm ist der Doppelpunkt S zu erkennen, der in verschiedener Weise vorher und nachher aufgelöst ist. Die Verhältnisse sind hier ganz analog, als wenn wir die kotierte Projektion bei einem hyperbolischen Paraboloid mit horizontal gelegener Tangentialebene im Scheitelpunkte S betrachten, wo die Umgebung des Scheitelpunktes S vor und nach dem Horizontalschnitt in den beiden Erzeugenden des Scheitelpunktes Hyperbeläste in dem einen oder anderen Paar Winkelräumen liefert (Fig. 10a, b). Ja, man kann, wenn man will, die Umgebung des Punktes S unserer Fläche $\mathfrak{F}$ geradezu als solchen Teil eines hyperbolischen Paraboloids ausgebildet denken.

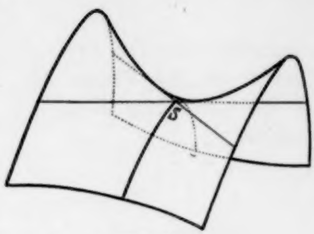

Fig. 10a.

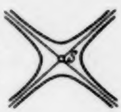

Fig. 10b.

Wenn man nun auch unsere so gewonnene Fläche sich in ihrer Gesamtheit vielleicht schwer vorzustellen vermag, so ist doch leicht ein anschauliches Bild von ihren einzelnen Teilen zu gewinnen, in die man sie sich zerlegt denken kann. Und zwar möge man sich außer dem Möbiusschen Blatt an der Hand der Fig. 9 jeden der vier Teile der Fläche $\mathfrak{F}$ vorstellen, die durch Verbiegen und Verzerren aus denjenigen vier Quadranten der Kreisfläche der Fig. 6 entstehen, wie sie durch die Durchmesser GI und KH geliefert werden. Den zugehörigen Halbmessern dieser Quadranten entsprechen ja die vier feingezeichneten Linien, die vom Punkte S des Schnittes 20 mm der Fig. 9 auslaufen. Man überzeugt sich dann vollends auch leicht, daß diese vier Teile der Fläche in der Tat zusammen der schlichten Kreisfläche der Fig. 6 entsprechen.

Um nun auch bequem die punktweise eineindeutige Abbildung der projektiven Ebene $\mathfrak{E}$ auf unsere Fläche $\mathfrak{F}$ der Fig. 9 überblicken zu können, wollen wir auf der Ebene $\mathfrak{E}$ in einfacher Weise ein Kurvennetz einzeichnen und dieses über die Halbkugel (Fig. 1 oder 7) und die Kreisfläche (Fig. 2) auf die Fläche $\mathfrak{F}$ übertragen. Zu dem Zweck führen wir ein rechtwinkliges Koordinatensystem mit dem Anfangspunkt M

und vertikaler z-Achse ein (Fig. 1). Die Gleichung der gegebenen Kugel sei:

$$(1) \qquad x^2 + y^2 + z^2 = r^2.$$

Die Hyperbeln der Fig. 6 und die zu ihnen deswegen noch hinzuzudenkenden Hyperbeln derselben Ebene, weil ja die Begrenzung der Fig. 6 kein größter Kugelkreis ist, liegen ferner auf den hyperbolischen Zylindern:

$$(2) \qquad y^2 - z^2 = c \qquad (-r^2 \leqq c \leqq +r^2),$$

entsprechen also den durch die beiden Gleichungen (1) und (2) gegebenen Kurven auf der Kugel. Die orthogonalen Projektionen dieser Kurven auf die horizontale Kreisfläche der Fig. 4 sind gegeben durch die Gleichungen:

$$(3) \qquad x^2 + 2y^2 = r^2 + c,$$

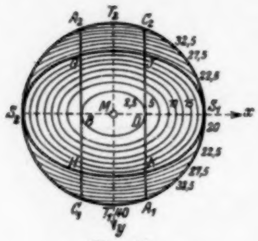

Fig. 11.

sind also Ellipsenbogen (Fig. 11). Für $c = 0$ erhält man eine Ellipse mit den Scheitelpunkten S_1, S_2, und zwar entspricht diese Ellipse den Durchmessern GI und HK der Fig. 6. Für $c = -r^2$ ergibt sich der Punkt M und für $c = +r^2$ ergeben sich die Schnittpunkte T_1, T_2 der y-Achse mit dem begrenzenden Kreise der Fig. 11.

Vom Mittelpunkte M der Kugel aus werden die Kurven auf der Halbkugel durch die Kegel

$$(4) \qquad cx^2 + (c - r^2)\cdot y^2 + (c + r^2)\cdot z^2 = 0$$

projiziert. Die Kurven der Halbkugel entsprechen also in der Ebene $\mathfrak{E}$, in die wir die vom Berührungspunkte M_0 ausgehenden ξ, η-Achsen parallel zu den x, y-Achsen einführen, den Kurven:

$$(5) \qquad c\cdot\xi^2 + (c - r^2)\cdot\eta^2 + (c + r^2)\cdot r^2 = 0 \qquad \text{(Fig. 12)}.$$

Für $c = 0$ ergeben sich also die Parallelgeraden $\eta = \pm r$ und für $-r^2 \leqq c < 0$ Ellipsen, für $0 < c \leqq r^2$ Hyperbeln, die sich für $\lim c = r^2$ auf den unendlich fernen Punkt T_0^∞ der η-Achse zusammenziehen. In der Fig. 12 ist außer diesen Kurven auch die den Parallelgeraden der Fig. 4 entsprechende Hyperbel eingezeichnet.

Es ist nun leicht zu übersehen, daß wir diese Kurvenschar in der Ebene $\mathfrak{E}$ unmittelbar den Höhenkurven unserer Fläche $\mathfrak{F}$ in der Fig. 9 zugeordnet annehmen können, so daß wir auch die zugehörigen Koten $\varkappa$ der Fig. 9 auf die Kurvenschar der Ebene $\mathfrak{E}$ übertragen können (vgl.

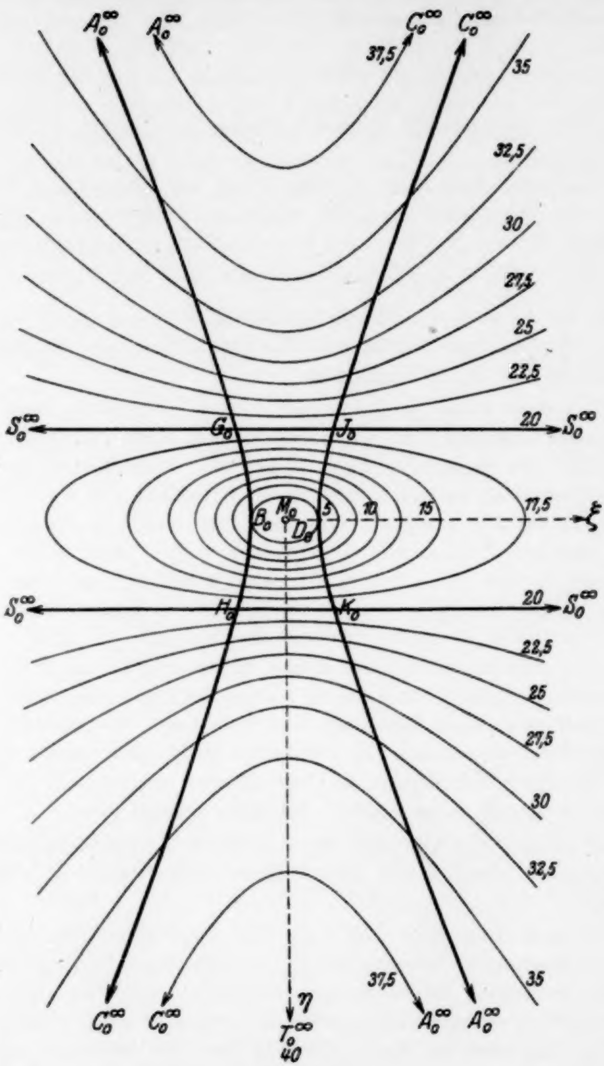
A_0^∞
A_0^∞
C_0^∞
C_0^∞
37,5
35
32,5
30
27,5
25
22,5
S_0^∞
G_0
J_0
20
S_0^∞
M_0
B_0
D_0
5
10
15
17,5
ξ
S_0^∞
H_0
K_0
20
S_0^∞
22,5
25
27,5
30
32,5
37,5
35
η
C_0^∞
C_0^∞
T_0^∞
40
A_0^∞
A_0^∞

Fig. 12.

Fig. 12). Wir wollen einfach folgende Beziehung zwischen den Parametern c und $\varkappa$ festsetzen:

$$(6) \qquad \varkappa = \frac{c+r^2}{2r^2}\cdot h \qquad \text{oder} \qquad (6\text{a}) \qquad c = r^2\left(\frac{2\varkappa}{h}-1\right),$$

wo h wieder die gesamte Höhe der Fläche, bei unserer Annahme also $h = 40$ mm, ist.

Sollen die Punkte B, D der Fig. 11 auf der Ellipse für $\varkappa = 2{,}5$ mm oder $c = -\frac{7}{8}r^2$ gelegen sein und analog die Punkte A_1, C_1, A_2, C_2 auf der Ellipse für $\varkappa = 37{,}5$ mm oder $c = +\frac{7}{8}r^2$, so muß ersichtlich der Abstand $MB = MD = \frac{r}{2\sqrt{2}}$ gewählt sein, wie es in den Figuren auch der Fall ist, für die noch $r_\varphi = 14$ mm gewählt ist. In den Fig. 6, 11 und 12 ist dann den einzelnen Kurven noch die Kote $\varkappa$ der ihnen entsprechenden Höhenkurven der Fig. 9 hinzugefügt.

Wir wollen weiter in der Ebene $\mathfrak{E}$ noch eine zweite Kurvenschar hinzunehmen, nämlich z. B. die konzentrischen Kreise mit dem Mittelpunkte M_0. Es ist einfach, diese Kreise auf die Halbkugel der Fig. 1, auf die Kreisfläche der Fig. 2 und auf die Kreisfläche der Fig. 6 zu übertragen. In der Fig. 2 ergeben sich dann die Kreise um M und in der Fig. 6 parallele Sehnen zum Durchmesser A_2C_1. Auch diese Kurven wollen wir je mit einer Indexziffer versehen denken, und zwar können wir als solche Indexziffer λ etwa den Radius der konzentrischen Kreise in der Fig. 2 oder 11 wählen, so daß die Indexziffer λ von 0 bis r variiert. Die Schnittpunkte dieser Kreise in der Fig. 11 mit der anderen Kurvenschar wollen wir auch auf die Kurven der Fig. 9 übertragen denken, so daß jede von diesen dann auch eine Skala trägt, der Indexziffer λ der Schnittpunkte entsprechend. Es hat keinen Zweck, dies, etwa auch durch Vervollständigung der Figuren, im einzelnen noch auszuführen, was ja gar keine Schwierigkeit bieten würde. Jedenfalls erkennt man:

Auf solche Weise läßt sich die punktweise eineindeutige Zuordnung der projektiven Ebene $\mathfrak{E}$ und unserer konstruierten Fläche $\mathfrak{F}$ vollständig übersehen.

Man kann dann auch jede Figur der projektiven Ebene $\mathfrak{E}$ auf die Fläche $\mathfrak{F}$ übertragen, insbesondere ein rechtwinkliges Koordinatensystem oder ein projektives Dreieckskoordinatensystem. Im Falle der Euklidischen Geometrie entspricht insbesondere der unendlichfernen Geraden von $\mathfrak{E}$, d. h. dem begrenzenden Kreise der Fig. 11, eine bestimmt angebbare Kurve der Fläche $\mathfrak{F}$, die auch durch die Punkte S, A, C und T der Höhenkurven 20, 37,5 und 40 mm der Fig. 9 hindurchgeht. Im Falle der hyperbolischen Maßbestimmung können wir etwa den begrenzenden Kreis

der Fig. 6 als Abbild des reellen absoluten Kegelschnittes ansehen und werden dann besser als die abzubildende projektive Ebene die Tangentialebene im Punkte S_2 der Fig. 7 (an Stelle der früheren Ebene $\mathfrak{E}$) wählen. Dann umfaßt der von den dünnen Linien der Fig. 9 gebildete Flächenteil, die Verbiegung der Kreisfläche der Fig. 6, noch abgesehen von der Begrenzung, die Gesamtheit der *eigentlichen* Punkte der hyperbolischen Ebene, während das Möbiussche Blatt einschließlich der Begrenzung die Gesamtheit aller *uneigentlichen* Punkte darstellt. Man kann also die Maßbestimmung in allen drei Arten der Geometrien unmittelbar auf unsere Fläche $\mathfrak{F}$ übertragen denken.

(Eingegangen am 8. 3. 1924.)

Bemerkungen über zentrische Kollineation[1]).

Von

Georg Heußel in Gießen.

Die folgenden Betrachtungen stehen im Zusammenhang mit dem Aufsatz des Herrn Pasch „Über zentrische Kollineation" in Bd. **90** der Annalen[2]). Zu dem dort behandelten ebenen Problem und zu ähnlichen führt ein Königsweg über die dritte Dimension. In Nr. 9 S. 106 jenes Aufsatzes wird auf einen Sonderfall folgender Aufgabe hingewiesen:

Von zwei Kegelschnitten c_1 und c_2 derselben Ebene Π ist eine gemeinsame Sehne s bekannt, es ist ein Schnittpunkt T gemeinsamer Tangenten zu suchen, so daß c_1 und c_2 zentrisch kollinear mit T als Zentrum und s als Achse sind. Ich behandle zunächst die umgekehrte Aufgabe und suche zu T die Gerade s. Vorausgesetzt sei, daß jede Gerade durch T sowohl mit c_1 als auch mit c_2 entweder zwei (reelle) oder gar keinen Punkt gemeinsam hat. Auf einer Geraden g, die mit Π nur T gemeinsam hat, nehmen wir zwei Punkte S und O an. S sei Spitze eines Kegels K_S mit der Leitlinie c_1. $p_1(p_2)$ sei die Polare von T in bezug auf c_1 (c_2). Die Ebene durch p_1 und S schneide die Ebene durch p_2 und O in der Geraden p. A_2 sei ein Punkt auf c_2, A_2O schneide K_S in A_3 und A_4. Die Ebene $E_3(E_4)$ durch A_3 (A_4) und p schneide K_S in dem Kegelschnitt c_3 (c_4) und Π in der Geraden s_3 (s_4). Dann sind c_1 und c_3 (c_4) zentrisch kollinear mit S als Zentrum und s_3 (s_4) als Achse. Diese räumliche Kollineation zwischen c_1 und c_3 (c_4) geht durch Projektion aus O auf Π über

---

[1]) Anm. d. Red.: Die Note von Pasch, Über zentrische Kollineation, Math. Ann. **90**, hat zwei Arbeiten hervorgerufen, die kurz nacheinander eingelaufen sind. Da sie, abgesehen von dem Ausgangspunkt, wenig Berührung miteinander haben, drucken wir sie im Einverständnis mit Herrn Pasch beide ab.

[2]) Da dort irrtümlich die Meinung geäußert ist, daß bei zentrischer Kollineation Zentrum und Achse auseinander liegen müssen, so hat Herr Pasch mich ermächtigt, diesen Irrtum, der einige Unrichtigkeiten im Verlauf der Betrachtungen zur Folge gehabt hat, hier zu berichtigen. (S. 107, Z. 3 von unten lies Pyrkosch.)

in eine ebene zentrische Kollineation zwischen c_1 und c_2 mit T als Zentrum und s_3 (s_4) als Achse.

Wir fassen zusammen: Sind c_1 und c_2 Kegelschnitte derselben Ebene Π und ist T ein Punkt, in dem die Involution konjugierter Strahlen in bezug auf c_1 und c_2 dieselbe ist, schneidet außerdem jede Gerade durch T entweder beide Kegelschnitte oder gar keinen, so gibt es auf jedem Kegel K_S mit der Leitlinie c_1 und der Spitze S zwei ebene Schnitte c_3 und c_4, die durch Projektion aus demselben Punkt O auf TS in c_2 übergehen; ihre Ebenen schneiden Π in s_3 und s_4, zwei gemeinsamen Sehnen von c_1 und c_2. Die Schnittpunkte von c_3 und c_4 mit Π liegen zugleich auf c_1 und c_2.

Um nun alle Fälle der gegenseitigen Lage von c_1 und c_2 zu erschöpfen, gehen wir von der räumlichen Kollineation zwischen c_1 und c_3 auf K_S aus und lassen c_2 entstehen, indem wir c_3 von O aus auf Π, die Ebene von c_1, projizieren. (Dabei fällt T jedesmal auf s_3, wenn O in der durch s_3 und S bestimmten Ebene liegt.)

I. c_1 und c_3 schneiden sich in zwei Punkten U und V.

1. O beliebig: c_1 und c_2 schneiden sich in U und V und im allgemeinen noch einmal in X und Y, den Schnittpunkten von c_4 mit Π. UV und XY sind Kollineationsachsen.

2. O auf K_S: c_1 und c_2 schneiden sich in U und V, c_4 wird (doppelt zu zählende) Mantellinie, X, Y, T fallen zusammen, c_1 und c_2 berühren sich zweipunktig, Zentrum ist der Berührungspunkt, Achse ist UV oder die Tangente des Berührungspunktes, in diesem Fall ist die Kollineation singulär.

3. O liege auf der Tangentialebene, die in U an K_S gelegt ist: c_1 und c_2 berühren sich in U und schneiden sich in V, die Tangentialebene berührt auch c_4, und der eine Schnittpunkt von c_4 mit Π, etwa X, fällt mit U zusammen. c_1 und c_2 berühren sich zweipunktig, c_1 und c_2 schneiden sich außerdem in V und Y. UV und UY sind Achsen.

4. O auf US. U, X, Y, T fallen zusammen, c_1 und c_2 berühren sich dreipunktig in U, schneiden sich in V, sind regulär kollinear in bezug auf UV und singulär kollinear in bezug auf die Tangente in U.

5. O liege auf der Schnittgeraden der beiden Ebenen, die K_S in U und V berühren, d. h. auf der Geraden, die zu UV in bezug auf K_S polar ist: X rückt mit U, V mit Y zusammen, c_1 und c_2 berühren sich doppelt, T liegt auf beiden Tangenten, die beiden Achsen fallen zusammen.

II. c_1 und c_3 berühren sich in U.

1. O beliebig: c_1 und c_2 berühren sich in U, schneiden sich in X und Y, die gemeinsame Tangente in U oder XY sind Achsen.

2. O auf K_S: c_1 und c_2 berühren sich doppelt, da X und Y zusammenfallen. T fällt auf X, die Kollineation ist regulär oder singulär, je nachdem die Tangente in U oder in X Achse ist.

3. O auf der Tangentialebene in U an K_S: c_1 und c_2 berühren sich dreipunktig in U und schneiden sich in Y. T liegt auf der Tangente durch U, entweder diese oder UY ist Achse.

4. O auf US: U, X, Y und T fallen zusammen, die Berührung wird vierpunktig, der Berührungspunkt Zentrum, die Tangente Achse einer regulären und einer singulären Kollineation.

Von den mannigfachen Zusammenhängen und Folgerungen, die sich aus diesen räumlichen Betrachtungen ergeben, greife ich nur eine ganz elementare heraus: Gesucht seien die gemeinsamen Sehnen zweier Kreise c_1 und c_2. Wir bestimmen einen Ähnlichkeitspunkt T im Endlichen und verfahren wie oben. O liege im Unendlichen, OT sei senkrecht Π; dann ist c_3 ein Kegelschnitt, dessen Ebene Π in der Potenzlinie s_3 von c_1 und c_2 schneidet, c_4 ein Kreis, dessen Ebene Π in der unendlich fernen Geraden s_4 schneidet. Sind die Kreise konzentrisch, so fallen s_3 und s_4 zusammen, es ist dann Fall I 5 verwirklicht.

(Eingegangen am 12. 11. 1923.)

Über zentrische Kollineation von Kegelschnitten.

Von

M. Frank in Simferopol (Krim Rußland).

In der Arbeit von M. Pasch: „Über zentrische Kollineation" (Math. Annalen **90**, 1/2, S. 103) steht: „Berühren die Kegelschnitte einander dreipunktig, so ist die Tangente des Berührungspunktes die eine jener beiden Schnittsehnen. Dies ist der einzige Fall, wo der Schnittpunkt gemeinsamer Tangenten auf einer zugeordneten Schnittsehne liegt" und weiter (S. 104) „Da Zentrum und Achse auseinanderliegen müssen, so kann ein Punkt mit Berührung höherer Ordnung zwischen K und K' nicht Zentrum, seine Tangente nicht Achse sein."

Beide Behauptungen, sowie alle Folgerungen aus ihnen sind irrtümlich. In der Tat, durch drei Paar zugeordnete Punkte A, B, C und A', B', C' ist eine zentrische Kollineation vollständig bestimmt, wobei fünf von ihnen willkürlich angenommen sein können, der sechste aber z. B. C' auf dem Strahle liegen muß, der durch das Kollineationszentrum P, das durch die Strahlen AA' und BB' bestimmt ist, und den Punkt C geht. Die Kollineationsachse wird aber nur durch die Lage des Punktes C' bestimmt. Nehmen wir also fünf willkürliche Punkte A, B, C, A', B' an, bestimmen das Kollineationszentrum P und einen Punkt G auf der Kollineationsachse, als Schnitt von AB und $A'B'$, so können wir immer die Gerade GP als Kollineationsachse annehmen, wobei der Punkt C' sich dann als Schnitt des Strahles PC mit der Geraden, die den Punkt F, Schnitt von GP mit AC (oder BC), mit A' (bzw. B') verbindet. Die Kollineation kann auch anders bestimmt werden, nämlich durch willkürliche Annahme von Zentrum und Achse, einem Paar zugeordneter Punkte A, A' auf einem Strahle, und einem Punkte B auf einem anderen. Daraus folgt, daß zu jedem Kegelschnitte K unendlich viele zentrisch kollineare Kegelschnitte K' konstruiert werden können und dabei so, daß die Kollineationsachse durch das Kollineationszentrum geht, wobei K und K' sich nicht notwendig berühren müssen.

Gerade diese Fälle zentrischer Kollineation, die M. Pasch für unmöglich hält, wollen wir etwas ausführlicher behandeln.

Nehmen wir bei gegebenem K das Kollineationszentrum P und die Kollineationsachse l ganz willkürlich an, nur so, daß P auf l liegt, ziehen durch P einen beliebigen Strahl, der K in A schneidet, und nehmen auf ihm den Punkt A' von K' an, so können alle anderen Punkte von K' konstruiert werden. Dabei wollen wir einstweilen P außerhalb von K annehmen. Man wird dann drei Fälle unterscheiden können.

1. Die Achse l geht durch P im Innern des Winkels der von den beiden Tangenten t und s aus P an K gebildet wird, dann schneidet l den Kegelschnitt K in zwei reellen Punkten N und M, die auch Schnittpunkte von K und K' sein müssen und durch die zwei Berührungspunkte T und S getrennt werden. Es müssen also dann außer N und M noch zwei reelle Schnittpunkte N' und M' der Kegelschnitte K und K' vorkommen. Die Kegelschnitte schneiden einander also in vier reellen Punkten. Die zweite, zum Zentrum P zugeordnete Achse l' ist $M'N'$.

Es ist bekannt (M. Pasch, l. c. S. 105), daß in diesem Falle im ganzen zwölf reelle verschiedene Kollineationen vorkommen. Sechs Schnitte von gemeinsamen Tangenten ergeben sechs Kollineationszentren, sechs gemeinsame Sehnen sechs Kollineationsachsen. Zu jedem von zwei Zentren, z. B. P und P' sind zwei Achsen l und l' und zu jeder von diesen Achsen dieselben zwei Zentren P und P' zugeordnet.

Wir wollen jetzt den Satz beweisen, *daß, falls eine Achse l durch das zugeordnete Zentrum P geht, auch die zweite zu P zugeordnete Achse l' durch das zweite, zu l zugeordnete Zentrum P' geht.* In der Tat, bilden die vier Schnittpunkte $MNM'N'$ ein gemeinsames, in die Kegelschnitte K und K' eingeschriebenes vollständiges Viereck, so bestimmen die drei Nebenpunkte des Vierecks E, F, G ein gemeinsames autopolares Dreieck von K und K'. Liegt der Punkt E im Innern der Kegelschnitte, so werden die anderen Punkte F und G auf der Geraden FG durch die Schnittpunkte mit den Seiten MN und $M'N'$ harmonisch geteilt. Andererseits bestimmen vier gemeinsame Tangenten s, t, s', t' ein Vierseit, dessen Nebenseiten e, f, g dasselbe autopolare Dreieck EFG bestimmen müssen, da zwei Kegelschnitte nur ein gemeinsames autopolares Dreieck haben können. Daraus folgt, daß die Zentren P und P' auf der Geraden FG (die wir auch e nennen können) liegen müssen. Die beiden anderen Nebenseiten g und f fallen mit EF und EG zusammen. Die Punkte P und P' werden durch G und F harmonisch geteilt. Daraus folgt, daß, falls eine von den Sehnen NM, die als zu P zugehörige Kollineationsachse angesehen werden kann, durch P geht, auch $N'M'$ durch P' durchgehen muß.

Eine Folgerung aus diesem Satz ist u. a., daß das Viereck $SS'TT'$, das durch vier Berührungspunkte der gemeinsamen Tangenten aus P gebildet wird, als Nebenecken die Zentren P, P' und den Punkt E hat, wobei die zugeordneten Achsen l und l' mit den Seiten PE und $P'E$ zusammenfallen. Dasselbe gilt von dem Viereck der Berührungspunkte gemeinsamer Tangenten aus P'.

2. Die Achse l läuft außerhalb des Winkels der beiden Tangenten, der die Kegelschnitte enthält. Dann schneidet l K nicht mehr. Dabei kann es vorkommen, daß

a) K und K' zwei reelle Schnittpunkte haben, durch welche die zweite zugeordnete Achse l' hindurchgeht.

b) K und K' sich berühren, und

c) keine gemeinsamen Punkte haben, aber in den letzten zwei Fällen muß jedenfalls der eine Kegelschnitt außerhalb des anderen liegen.

Der für den Fall 1 bewiesene Satz, daß auch die zweite zugeordnete Achse durch das zweite Zentrum geht, gilt auch hier, trotzdem in diesem Falle kein reelles gemeinschaftliches autopolares Dreieck existiert. Es ist aber auch hier der Pol E und die gemeinsame Polare PP' vorhanden, und die als Folgerung zu dem Satze gegebene Konstruktion mittels des Vierecks $SS'TT'$ ergibt die Lage von P' und E, also auch von l'.

3. Fällt die Achse l mit einer der Tangenten von P an K zusammen, so können wir diesen Fall als Grenzfall ansehen, wenn l, um den Punkt P sich drehend, aus dem Inneren des Tangentenwinkels in den Nebenwinkel übergeht. Die Schnittpunkte M und N fallen beide mit dem Berührungspunkte T zusammen, aber noch ein dritter Schnittpunkt M' oder N' muß dann auch mit T zusammenfallen. Die Kegelschnitte K und K' berühren sich dreipunktig. Die zweite zugeordnete Achse ist dann die Schnittsehne, das zweite Zentrum P' ist der Berührungspunkt T beider Kegelschnitte. Zum Berührungspunkte T als Zentrum ist aber nur eine Achse — die Schnittsehne — zugeordnet, weil jeder Strahl durch T jeden Kegelschnitt nur noch in einem Punkte schneidet.

Es bleiben noch die Fälle zu untersuchen, wo das Zentrum P im Innern von K oder auf K liegt.

Diese Fälle sind aber, einen ausgenommen, in den oben behandelten enthalten, weil im Falle 2a das zweite Zentrum P' im Innern von K liegen muß, und im Falle 2b sowie 3 auf K selbst. Es ist nur noch ein Fall übrig, wenn P und P' zusammenfallen, und das kommt vor, wenn K und K' sich vierpunktig berühren. Die gemeinsame Tangente ist dann die Kollineationsachse.

Wir können somit auf Grund des Vorangehenden behaupten, daß die Kollineationsachse durch das Zentrum hindurchgehen kann, bei verschiedenster Lage der Kegelschnitte, nur den Fall ausgenommen, wo der eine Kegelschnitt vollständig im Innern des anderen liegt und die Kegelschnitte dabei keine Berührung höherer Ordnung miteinander haben. Im Falle doppelter Berührung beider Kegelschnitte erhalten wir zwar mittels zweier Tangenten und einer Sehne drei Achsen und drei Zentren, die auf den Achsen liegen und ein Dreieck bilden, aber jedem Zentrum ist die *gegenüberliegende* Seite des Dreiecks als Achse zugeordnet[1]). Die behandelten speziellen Fälle zentrischer Kollineation führen zur Lösung vieler Aufgaben, von denen wir nur auf folgende aufmerksam machen.

Aufgabe I. *Zu einem Kegelschnitte K in einem gegebenen Punkte P den Krümmungskreis zu konstruieren.*

Der Krümmungskreis, der mit dem Kegelschnitte eine dreipunktige Berührung hat, steht mit ihm in zentrischer Kollineation, wobei als Zentrum der Berührungspunkt P erscheint. Zeichnen wir einen beliebigen Kreis K_1, der im Punkte P den Kegelschnitt einfach berührt, und bestimmen die Kollineationsachse l_1 von K und K_1, so ist leicht zu beweisen, daß für einen anderen Kreis K_2, der ebenso den Kegelschnitt in P berührt, die zugehörige Kollineationsachse l_2 zu l_1 parallel sein muß. Nehmen wir also l' parallel zu l_1 durch den Punkt P, so ist das die gesuchte Achse für den Krümmungskreis und Kegelschnitt. Der Durchschnitt von l' und K gibt einen zweiten Punkt des Kreises und damit ist die Aufgabe gelöst.

Falls vom Kegelschnitt K nur fünf Punkte gegeben sind, konstruieren wir im Punkte P die Tangente und Normale, nehmen ebenso einen Kreis durch P mit Zentrum auf der Normale, bestimmen mittels dreier anderer Punkte von K die Achse l_1 und ziehen dann zu ihr die Parallele l' durch P. Es bleibt dann noch ein weiterer Punkt des Krümmungskreises K zu bestimmen.

Aufgabe II. *Einen Kegelschnitt zu konstruieren, der zentrisch kollinear dem Krümmungskreise in seinem Scheitel ist.*

Ziehen wir die Tangente l im Punkte P des Krümmungskreises und einen Durchmesser PA, der den Kreis im Punkte A zum zweiten Male schneidet, und nehmen wir auf PA den zu A zugehörigen Punkt A' des gesuchten Kegelschnitts an, so bestimmen sich alle anderen Punkte von K mittels einfacher Konstruktion: Schneidet ein beliebiger Strahl g

[1]) Bei M. Pasch (l. c. S. 105) steht irrtümlich, daß in diesem Falle vier Kollineationen vorkommen.

durch P den Kreis in B und ziehen wir die Gerade AB bis zum Schnitte C mit der Tangente l, so erhalten wir den Punkt B' als Schnitt von CA' und g.

PA' ist eine der Hauptachsen des Kegelschnittes. Liegt der Punkt A' mit dem Kreise K auf der gleichen Seite der Tangente l, so erhalten wir eine Ellipse, wobei, falls A' innerhalb des Kreises liegt, PA' die kleine Achse ist, falls außerhalb, die große. Liegen A' und K auf verschiedenen Seiten von l, so erhalten wir eine Hyperbel. Ist A' im Unendlichen, so erhalten wir eine Parabel.

(Eingegangen am 16. 12. 1924.)

Bemerkung über die ebenen Elementarkurven 3. Ordnung.

Von

Otto Haupt in Erlangen.

1. *Einleitung.* Unter einer *„völlig stetigen“* (ebenen) Kurve versteht man nach Herrn Juel[1]) eine stückweise stetig differenzierbare Kurve, also eine (im projektiven Sinne geschlossene) Kurve, für welche, mit Ausnahme höchstens endlich vieler Stellen, an jeder Stelle genau *eine* vordere und *eine* hintere Halbtangente existiert, die entgegengesetzt gerichtet sind und sich stetig mit der Stelle ändern[2]). Kurven, die sich aus einer endlichen Anzahl von konvexen Bogen („Elementarbogen“[3])) zusammensetzen, heißen *„Elementarkurven“*[1]).

Herr Juel hat nun den Satz bewiesen[4]): *Die völlig stetigen, ebenen Kurven* 3. *Ordnung sind Elementarkurven.*

Im nachstehenden soll gezeigt werden, daß die Voraussetzung der völligen Stetigkeit entbehrt werden kann, daß also folgende Behauptung richtig ist: *Die ebenen Kurven* 3. *Ordnung sind Elementarkurven.*

Nachdem man einmal weiß, daß die Kurven 3. Ordnung stets Ele-

---

[1]) a) C. Juel, Einleitung in die Theorie der ebenen Elementarkurven 3. und 4. Ordnung (D. Kgl. Danske Vidensk. Selsk. Skrifter, 7. R., Naturvidensk. og Mathem. Afd., 11, 2, Kopenhagen 1914, S. 113 und 125);

b) vgl. auch C. Juel, Einige Sätze über ein- und mehrteilige Elementarkurven 4. Ordnung (Math. Ann. 76 (1915), S. 343–346, bes. § 2, S. 345).

[2]) Wegen der Begriffe: Kurve, stetig differenzierbare Kurve, Halbtangente usw. siehe A. Rosenthal, Über die Singularitäten der reellen ebenen Kurven (Math. Ann. 73 (1913), S. 481 ff.).

[3]) Juel, l. c. [1]), a) S. 116. Bezüglich der Definition des konvexen Bogens vgl. auch Nr. 3 der vorliegenden Arbeit. Die von uns benutzte Definition ist gleichbedeutend mit der von Herrn Juel zugrunde gelegten.

[4]) Juel, l. c. [1]), a) S. 136 ff. Auch mehrteilige Kurven können zugelassen werden. Der Satz gilt, wie Herr Juel (l. c., S. 114) hervorhebt, nicht mehr für Kurven 4. Ordnung.

mentarkurven sind, überträgt sich die von Herrn Juel[5]) (für völlig stetige Kurven) entwickelte Theorie im wesentlichen unverändert auf die allgemeinen Kurven 3. Ordnung.

2. Weil jeder Teilbogen eines konvexen Bogens wieder konvex ist, folgt — etwa mit Hilfe des Heine-Borelschen Überdeckungssatzes — zunächst[6]): Damit die (ebene) Kurve $\mathfrak{C}$ eine Elementarkurve sei, ist notwendig und hinreichend, daß zu jeder Stelle P von $\mathfrak{C}$ eine Umgebung[7]) auf $\mathfrak{C}$ gehört, die in endlich viele Konvexbögen zerlegbar ist.

Zum Beweise unserer Behauptung genügt es deshalb, *Bogen* 3. Ordnung zu betrachten.

Diese Bogen können wir überdies als (*beschränkte*) *Jordanbogen* voraussetzen. Bei der Bestimmung der „Ordnung" einer Kurve wird nämlich jeder *mehrfache Punkt* der Kurve mit seiner Vielfachheit in Anschlag gebracht[8]); eine Kurve 3. Ordnung kann demgemäß höchstens *einen* mehrfachen Punkt vom Vielfachheitsgrad 2 besitzen, sie ist also in endlich viele Bogen zerlegbar, von denen keiner einen mehrfachen Punkt besitzt. Und wir haben nur *beschränkte* derartige Bogen in Betracht zu ziehen, weil die Umgebung[9]) eines unendlich fernen Punktes auf $\mathfrak{C}$ (bzw. ein konvexer Bogen, der einen unendlich fernen Punkt enthält) definiert ist als das projektive Bild eines beschränkten (bzw. eines beschränkten konvexen) Bogens[8]).

Wir können uns sogar *damit begnügen, die Umgebung eines Endpunktes*[9]) P_0 *von Jordanbogen 3. Ordnung zu betrachten.* Denn die Umgebung $\mathfrak{U}$ eines *inneren*[9]) Punktes P_0 auf einem Jordanbogen $\mathfrak{B}$ ist die Vereinigung der beiden einseitigen Umgebungen $\mathfrak{U}^+$ und $\mathfrak{U}^-$ von P_0 auf $\mathfrak{B}$. Also ist $\mathfrak{U}$ in endlich viele Konvexbogen zerlegbar, wenn dies für $\mathfrak{U}^+$ sowohl als auch für $\mathfrak{U}^-$ zutrifft. Ohne Einschränkung der Allgemeinheit dürfen wir schließlich *voraussetzen, daß der betrachtete Bogen keine den Endpunkt* P_0 *enthaltende Strecke als Teilbogen aufweist*; denn eine solche Strecke wäre bereits eine konvexe Umgebung von P_0 auf dem Bogen.

---

[5]) Juel, l. c. [1]), a) S. 137 ff.

[6]) Juel, l. c. [1]), a) S. 126.

[7]) Wegen des hier und im folgenden verwendeten Begriffes: Umgebung eines Punktes auf einer Kurve sei verwiesen auf die Arbeit „Über Kurven endlicher Ordnung" (Math. Zeitschr. 19 (1924), § 2, S. 288, Nr. 9). Im folgenden zitiert mit A.

[8]) Rosenthal, l. c. [2]), S. 487.

[9]) A., Nr. 9.

3. Ehe wir zum Beweise des in Nr. 1 aufgestellten Satzes, in der aus Nr. 2 sich ergebenden Fassung, übergehen, sei an folgendes erinnert[10]): Es bedeute $\mathfrak{B}$ einen Jordanbogen mit den (voneinander verschiedenen) Endpunkten P_0 und P_1. Dabei kann $\mathfrak{B}$ selbst eine Strecke sein oder $\mathfrak{B}$ kann Strecken als Teilbogen enthalten. *Wir bezeichnen mit* $K_{\mathfrak{B}}$ *die (abgeschlossene) konvexe Hülle von* $\mathfrak{B}$, *mit* $\mathfrak{G}(K_{\mathfrak{B}})$ *die Begrenzung von* $K_{\mathfrak{B}}$. *Ist* $\mathfrak{B}$ *Teilmenge von* $\mathfrak{G}(K_{\mathfrak{B}})$ *oder auch mit* $\mathfrak{G}(K_{\mathfrak{B}})$ *identisch, so heißt* $\mathfrak{B}$ *konvex.*

Im Falle $\mathfrak{B}$ ein beliebiger (nicht notwendig konvexer) Jordanbogen ist, besteht $\mathfrak{G}(K_{\mathfrak{B}})$ aus *Punkten der beiden folgenden Arten*:

I. Der Punkt B (Punkt erster Art) von $\mathfrak{G}(K_{\mathfrak{B}})$ gehört zu $\mathfrak{B}$. Falls $\mathfrak{B}$ keine Strecke ist, gibt es immer mindestens drei verschiedene derartige Punkte, die nicht der gleichen Geraden angehören.

II. Der Punkt $\mathsf{\Gamma}$ (Punkt zweiter Art) von $\mathfrak{G}(K_{\mathfrak{B}})$ gehört nicht zu $\mathfrak{B}$, sondern ist innerer Punkt der abgeschlossenen Verbindungsstrecke $\overline{\mathsf{B}\mathsf{B}'}$ zweier Punkte B und B' erster Art von $\mathfrak{G}(K_{\mathfrak{B}})$. Jeder Punkt im Innern der Strecke $\overline{\mathsf{B}\mathsf{B}'}$ ist ebenfalls ein Punkt $\mathsf{\Gamma}$ zweiter Art. Die Gerade β, welche B und B' verbindet, ist Stützgerade an $K_{\mathfrak{B}}$ längs der Strecke $\overline{\mathsf{B}\mathsf{B}'}$. Durch $\mathsf{\Gamma}$ sind B und B' eindeutig bestimmt.

Es gilt der, für das Folgende nützliche, *Hilfssatz*: Es sei $\mathfrak{B}$ *nichtkonvex*, es bezeichne $\mathsf{\Gamma}$ einen Punkt zweiter Art auf $\mathfrak{G}(K_{\mathfrak{B}})$, ferner B und B' die beiden, durch $\mathsf{\Gamma}$ bestimmten, Punkte erster Art auf $\mathfrak{G}(K_{\mathfrak{B}})$ und β ihre Verbindungsgerade. Dann gibt es auf $\mathfrak{G}(K_{\mathfrak{B}})$ immer Punkte $\mathsf{\Gamma}$ der folgenden Beschaffenheit: Mindestens einer der Punkte B und B' ist weder ein Endpunkt von $\mathfrak{B}$ noch Endpunkt einer auf β liegenden, nur Punkte von $\mathfrak{B}$ enthaltenden, abgeschlossenen Strecke, deren zweiter Endpunkt mit einem Endpunkt von $\mathfrak{B}$ zusammenfällt.

Der Beweis des Hilfssatzes wird in Nr. 7 nachgetragen.

4. Wir wenden uns zum Beweise unserer Behauptung, daß eine hinreichend kleine Umgebung des Endpunktes eines Jordanbogens 3. Ordnung konvex ist. Dieser Beweis soll indirekt geführt werden.

Der zu betrachtende Jordanbogen $\mathfrak{B}$ von 3. Ordnung sei umkehrbar eindeutiges und stetiges Abbild der Strecke $0 \leqq t \leqq 1$; dabei entspreche P_0 dem Parameterwerte $t_0 = 0$, in Zeichen: $P_0\{t_0\}$. Die zu untersuchende *Umgebung* $\mathfrak{U}(\varepsilon)$ $\{t_0 \leqq t \leqq t_0 + \varepsilon = t_1\}$ *von* P_0 *auf* $\mathfrak{B}$ mit dem zweiten

[10]) Siehe etwa: J. Hjelmslev, Contribution à la géométrie infinitésimale de la courbe réelle (Oversigt over det Kgl. Danske Vidensk. Selsk. Forh. 1911, Nr. 5, § 1, S. 434—442).

Endpunkte $P_1\{t_1\}$ sei *so gewählt*[11]), daß *erstens* $\mathfrak{U}(\varepsilon)$ *von* P_0 *aus unter einem Winkel* $\varphi < \frac{\pi}{2}$ *erscheint* und daß *zweitens* $\mathfrak{U}(\varepsilon)$ *relativ einfach ist bezüglich* P_0; letzteres soll besagen[12]), daß $\mathfrak{U}(\varepsilon)$ mit jedem von P_0 ausgehenden Halbstrahle s höchstens einen, von P_0 verschiedenen gewöhnlichen Punkt oder eine, P_0 nicht enthaltende, Strecke (d. h. einen „von P_0 verschiedenen, *auf s verlängerten*[12]) *Punkt von* $\mathfrak{B}$") gemeinsam hat.

Diese Wahl von $\mathfrak{U}(\varepsilon)$ ist möglich: Es gibt ein $\eta_0 > 0$, so daß für $0 < \varepsilon \leqq \eta_0$ alle $\mathfrak{U}(\varepsilon)$ diese Eigenschaften besitzen. Denn erstens ist jeder Bogen endlicher Ordnung überall vorn und hinten differenzierbar[13]), so daß eine genügend kleine, einseitige Umgebung jedes Punktes von eben diesem Punkte aus unter beliebig kleinem Winkel erscheint; zweitens ist jeder Bogen 3. Ordnung in endlich viele, relativ einfache Bogen zerlegbar[14]).

Es sei von jetzt ab *immer* $0 < \varepsilon \leqq \eta_0$ *vorausgesetzt*. Jeder auf einem Halbstrahl s durch P_0 gelegene innere (gewöhnliche oder auf s verlängerte) Punkt P von $\mathfrak{U}(\varepsilon)$ ist dann *Schnitt*punkt auf s [15]).

Ist $\mathfrak{U}(\varepsilon)$ *nicht konvex*, so gibt es nach dem Hilfssatze der Nr. 3 mindestens zwei voneinander verschiedene gewöhnliche Punkte $\mathsf{B}\{t(\varepsilon)\}$ und $\mathsf{B}'\{t'(\varepsilon)\}$ von $\mathfrak{U}(\varepsilon)$, deren Verbindungsstrecke $\overline{\mathsf{B}\mathsf{B}'}$ zwar zu $\mathfrak{G}(K_{\mathfrak{U}(\varepsilon)})$ nicht aber zu $\mathfrak{U}(\varepsilon)$ gehört und wobei B und B' nicht gleichzeitig mit den beiden (gewöhnlichen oder auf der Verbindungsgeraden β von B und B' verlängerten) Endpunkten von $\mathfrak{U}(\varepsilon)$ zusammenfallen[16]). Dabei sei etwa $t_0 \leqq t(\varepsilon) < t'(\varepsilon) \leqq t_1 = t_0 + \varepsilon$.

Unter der *weiteren Annahme*, daß $\mathfrak{U}(\varepsilon)$ *von* 3. *Ordnung*[17]) sei, werden wir jetzt (Nr. 5) zeigen: Der eben genannte Punkt B' muß für *jedes* ε $(0 < \varepsilon \leqq \eta_0)$ zusammenfallen mit dem (von P_0 verschiedenen) Endpunkte P_1 von $\mathfrak{U}(\varepsilon)$, der evtl. auf β verlängert ist; dabei ist, um es nochmals hervorzuheben, $\mathfrak{U}(\varepsilon)$ als nicht-konvex angenommen. Und von da aus werden wir schließlich (Nr. 6) zu einem Widerspruch gelangen.

---

[11]) Dabei ist $\varepsilon > 0$ eine reelle, beliebig klein gewählte Zahl. $\mathfrak{U}(\varepsilon)$ ist nach den in Nr. 2 gemachten Voraussetzungen keine Strecke.

[12]) A., Nr. 6.

[13]) A., Nr. 21, Fußnote [21]).

[14]) A., Nr. 19.

[15]) A., Nr. 16. — Definition von „Schnittpunkt" und „Stützpunkt" a. a. O. (Nr. 12).

[16]) Die Ausdrucksweise: „B (bzw. B') fällt mit einem *auf* β *verlängerten* inneren Punkte (bzw. Endpunkt) von $\mathfrak{U}(\varepsilon)$ zusammen" soll besagen, daß z. B. B einer auf β liegenden abgeschlossenen, nur aus Punkten von $\mathfrak{U}(\varepsilon)$ bestehenden Strecke angehört, die nur innere Punkte (bzw. die einen Endpunkt) von $\mathfrak{U}(\varepsilon)$ enthält.

[17]) Wir nehmen an, daß $\mathfrak{U}(\varepsilon)$ von 3. Ordnung sei, weil jeder Bogen von 2. (bzw. 1.) Ordnung konvex ist.

5. Zunächst können die eben (Nr. 4) erwähnten Punkte B und B' nicht beide mit *inneren* (gewöhnlichen oder auf β verlängerten) Punkten von $\mathfrak{U}(\varepsilon)$ zusammenfallen. Denn andernfalls ist β (die Verbindungsgerade von B und B') Stützgerade von $\mathfrak{U}(\varepsilon)$ mit den *Stützpunkten*[18]) B und B'; folglich[15]) hätte eine zu β hinreichend benachbarte Parallele mit $\mathfrak{U}(\varepsilon)$ mindestens vier verschiedene (gewöhnliche oder verlängerte[19])) Punkte gemeinsam, während doch $\mathfrak{U}(\varepsilon)$ von der Ordnung 3 sein soll.

Ferner können die in Rede stehenden Punkte B und B' auch nicht mit P_0 zusammenfallen. Für B' ist dies wegen der oben getroffenen Festsetzung $t_0 \leqq t < t' \leqq t_1$ von selbst ausgeschlossen. Nimmt man aber an, es falle B mit P_0 zusammen, so würde (vgl. Nr. 4) der durch B mitbestimmte Punkt B' mit einem *inneren* (gewöhnlichen oder auf β verlängerten) Punkt S von $\mathfrak{U}(\varepsilon)$ zusammenfallen; einer der von P_0 ausgehenden, auf β liegenden, Halbstrahlen hätte also mit $\mathfrak{U}(\varepsilon)$ den *Stützpunkt* S gemeinsam, während doch (gemäß Nr. 4) $\mathfrak{U}(\varepsilon)$ als relativ einfach bezüglich P_0 vorausgesetzt war. P_0 kann übrigens nach den (in Nr. 2) gemachten Annahmen auf keiner Geraden verlängerter Endpunkt von $\mathfrak{U}(\varepsilon)$ sein.

Falls $\mathfrak{U}(\varepsilon)$ nicht-konvex, ist es also, wie behauptet, immer möglich, durch den Endpunkt P_1 von $\mathfrak{U}(\varepsilon)$ eine Gerade β zu legen, welche $\mathfrak{U}(\varepsilon)$ in einem, von P_1 verschiedenen, inneren Punkte B stützt. P_1 und B können dabei auf β verlängerte Punkte von $\mathfrak{U}(\varepsilon)$ sein[19]). Außer P_1 und B hat β keine Punkte mit $\mathfrak{U}(\varepsilon)$ gemeinsam, weil $\mathfrak{U}(\varepsilon)$ von 3. Ordnung sein soll und weil B Stützpunkt ist[15]).

6. Es sei jetzt $\varepsilon = \varepsilon_0$ fest gewählt gemäß der Bedingung $0 < \varepsilon_0 \leqq \eta_0$ und es werde wieder $\mathfrak{U}_0 = \mathfrak{U}(\varepsilon_0)$ als nicht-konvex angenommen. Es bezeichne weiter β_0 die (zufolge Nr. 5 vorhandene) durch den Endpunkt P_1 von $\mathfrak{U}_0$ laufende Gerade, welche $\mathfrak{U}_0$ in dem von P_1 verschiedenen inneren Punkte B_0 von $\mathfrak{U}_0$ stützt. Schließlich bedeute $\mathfrak{H}_0^+$ diejenige der beiden, von β_0 gebildeten offenen Halbebenen, in der, weil β_0 Stützgerade von K_{u}, ist, P_0 und überhaupt alle Punkte von $\mathfrak{U}_0$ liegen, mit Ausnahme der (evtl. auf β_0 verlängerten) Punkte P_1 und B_0.

Sei *zunächst* $\mathsf{B}_0\{t(\varepsilon_0)\}$ *gewöhnlicher Punkt von* $\mathfrak{U}_0$ *auf* β_0. Der von B_0 ausgehende, in $\mathfrak{H}_0^+$ verlaufende Halbstrahl s_0 durch P_0 zerlegt $\mathfrak{H}_0^+$ in zwei (offene) Winkelgebiete: $\sphericalangle[P_1\mathsf{B}_0P_0]$ und $\sphericalangle[P_0\mathsf{B}_0R_0]$; dabei bedeutet

---

18) A., Nr. 12. Fällt B oder B' mit einem auf β verlängerten Punkt von $\mathfrak{U}(\varepsilon)$ zusammen, so ist unter dem „Stützpunkt" eben dieser verlängerte Punkt zu verstehen. — Die Bezeichnungen der Nr. 4 des Textes sind auch in gegenwärtiger Nr. 5 beibehalten.

19) Zwei verlängerte Punkte eines Bogens auf der gleichen Geraden heißen „verschieden", wenn die sie bildenden, abgeschlossenen Strecken punktfremd sind.

R_0 einen von B_0 verschiedenen Punkt des auf β_0 liegenden, von B_0 ausgehenden, P_1 nicht enthaltenden Halbstrahles. Die den Intervallen $t_0 \leqq t \leqq t(\varepsilon_0)$ bzw. $t(\varepsilon_0) \leqq t \leqq t_1$ entsprechenden Teilbogen $\widehat{P_0\mathsf{B}_0}$ bzw. $\widehat{\mathsf{B}_0P_1}$ von $\mathfrak{U}_0$ verlaufen, abgesehen von den Endpunkten P_0 und B_0 bzw. B_0 und (dem evtl. auf β_0 verlängerten Endpunkte) P_1, ganz in $\sphericalangle[P_0\mathsf{B}_0R_0]$ bzw. in $\sphericalangle[P_1\mathsf{B}_0P_0]$ [20]). Denn $\mathfrak{U}_0$ ist (Nr. 4) relativ einfach bezüglich P_0, kann also mit dem von B_0 ausgehenden Halbstrahl s_0, außer P_0 und B_0, höchstens *einen* weiteren Punkt gemeinsam haben, der überdies *nicht* auf der Strecke $\overline{\mathsf{B}_0P_0}$ liegen darf. Dies ist aber ausgeschlossen, weil $\mathfrak{U}_0$ von P_0 aus unter einem Winkel $\varphi < \frac{\pi}{2}$ erscheint (Nr. 4). B_0 ist überdies Schnittpunkt von $\mathfrak{U}_0$ auf s_0.

Wäre nun auch $\widehat{P_0\mathsf{B}_0}$, d. h. $\mathfrak{U}_1 = \mathfrak{U}(\varepsilon_1)\{t_0 \leqq t \leqq t(\varepsilon_0) = t_0 + \varepsilon_1\}$ $(0 < \varepsilon_1 < \varepsilon_0)$, nicht konvex, so ließe sich durch B_0 die Gerade β_1 legen, welche $\mathfrak{U}_1$ in dem inneren (evtl. auf β_1 verlängerten) Punkte B_1 stützt. B_1 ist verschieden von P_0 und von dem (evtl. auf β_1 verlängerten) Endpunkte von $\mathfrak{U}_1$, mit welchem B_0 zusammenfällt. Folglich ist B_1 innerer Punkt von $\sphericalangle[P_0\mathsf{B}_0R_0]$ und P_0 ist daher (innerer) Punkt des Winkelgebietes $\sphericalangle[P_1\mathsf{B}_0\mathsf{B}_1]$. Auf β_1 muß B_0 Schnittpunkt von $\mathfrak{U}_0$ sein oder mit einem auf β_1 verlängerten derartigen Schnittpunkte zusammenfallen [21]); denn sonst würde B_0 mit einem Stützpunkte von $\mathfrak{U}_0$ auf β_1 zusammenfallen und dann hätte eine geeignet gewählte, zu β_1 benachbarte Gerade mit $\mathfrak{U}_0$ mindestens vier verschiedene Punkte gemeinsam — einerlei ob die Umgebungen von B_1 und B_0 auf $\mathfrak{U}_0$ beide in der gleichen, durch β_1 bestimmten Halbebene verlaufen oder in verschiedenen Halbebenen [15]).

Es tritt also $\mathfrak{U}_0$ im Punkte B_0 aus dem Gebiete $\sphericalangle[P_1\mathsf{B}_0\mathsf{B}_1]$ über in das Gebiet $\sphericalangle[\mathsf{B}_1\mathsf{B}_0R_0]$. Da P_0 in $\sphericalangle[P_1\mathsf{B}_0\mathsf{B}_1]$ liegt, so muß $\mathfrak{U}_0$, außer B_0, noch mindestens einen weiteren *Schnittpunkt* auf β_1 besitzen [22]). Da andererseits β_1 mit $\mathfrak{U}_0$ außer dem Schnittpunkt B_0 und dem Stützpunkt B_1 weitere Punkte nicht gemeinsam haben kann (Nr. 5), sind wir bei einem Widerspruche angelangt.

Die oben gemachte Einschränkung, daß B_0 gewöhnlicher Punkt von $\mathfrak{U}_0$ auf β_0 sei, sollte nur eine bequemere Formulierung unserer Überlegungen ermöglichen. Diese lassen sich unverändert durchführen auch *im Falle eines auf β_0 verlängerten Punktes B_0 von $\mathfrak{U}_0$*.

---

[20]) B_0 kann nicht mit einem auf s_0 verlängerten Punkte von $\mathfrak{U}_0$ zusammenfallen. Denn sonst würde B_0 auf β_1 (siehe im Text weiter unten) einen Stützpunkt von $\mathfrak{U}_0$ liefern. Vgl. auch [21]).

[21]) Das wäre nicht der Fall, wenn B_0 mit einem auf s_0 verlängerten Punkte von $\mathfrak{U}_0$ zusammenfiele.

[22]) A., Nr. 13.

7. Nachtrag zu Nr. 3. Der *Beweis des Hilfssatzes* der Nr. 3 läßt sich etwa so führen: Es gibt Punkte Γ zweiter Art auf $\mathfrak{G}(K_{\mathfrak{B}})$. Denn andernfalls wären *alle* Punkte der einfachen geschlossenen Jordankurve $\mathfrak{G}(K_{\mathfrak{B}})$ Punkte B erster Art; diese geschlossene Kurve wäre also Teilmenge des Bogens $\mathfrak{B}$, was unmöglich ist.

Es seien P_0 und P_1 die Endpunkte von $\mathfrak{B}$. Angenommen nun, der Hilfssatz wäre nicht richtig, dann müßte die Gesamtheit der Punkte Γ von $\mathfrak{G}(K_{\mathfrak{B}})$ das Innere einer *einzigen* Strecke $\overline{P_0' P_1'}$ bilden, welche Teilmenge der Strecke $\overline{P_0 P_1}$ oder (wenn P_0 mit P_0' und P_1 mit P_1' zusammenfällt) mit ihr identisch ist; die abgeschlossenen Strecken $\overline{P_0 P_0'}$ und $\overline{P_1 P_1'}$ gehören zu $\mathfrak{G}(K_{\mathfrak{B}})$.[23]) Die nicht zum Innern von $\overline{P_0' P_1'}$ gehörigen Punkte von $\mathfrak{G}(K_{\mathfrak{B}})$ wären somit *sämtlich* von erster Art und folglich der von ihnen gebildete Jordanbogen $\mathfrak{B}^*$ mit den Endpunkten P_0' und P_1' Teilmenge von $\mathfrak{B}$. Da andererseits aber P_0 und P_1 innere oder Endpunkte von $\mathfrak{B}^*$ sind, so muß $\mathfrak{B}$ mit $\mathfrak{B}^*$ identisch sein; d. h. $\mathfrak{B}$ ist konvex gegen die Voraussetzung.

Erlangen, Math. Seminar d. Univ., 1. XII. 1923.

[23]) Dabei soll P_0' auf $\overline{P_0 P_1}$ zwischen P_0 und P_1' liegen.

(Eingegangen am 8. 12. 1923.)

Über Greensche Randbedingungen.

Von

O. Haupt in Erlangen und E. Hilb in Würzburg.

§ 1.

Fragestellung.

1. Es bedeute $\mathfrak{K}$ die Peripherie des Einheitskreises in der Ebene der rechtwinkligen kartesischen Koordinaten x, y. Die Bogenlänge auf $\mathfrak{K}$ sei mit s bezeichnet $(0 \leqq s \leqq 2\pi)$, das Innere $x^2 + y^2 < 1$ des Einheitskreises mit $\mathfrak{K}^*$.

Eine bekannte *Randwertaufgabe* besteht dann in der Bestimmung aller Werte des Parameters λ, für welche die Differentialgleichung $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \lambda u = 0$ eine in $\mathfrak{K} + \mathfrak{K}^*$ stetig differenzierbare, eindeutige und *auf* $\mathfrak{K}$ *verschwindende Lösung* $u(x, y)$ besitzt. Wie man weiß, gibt es abzählbar unendlich viele, ausnahmslos reelle derartige Werte λ und dementsprechend abzählbar unendlich viele Lösungen $u_j(x, y)$, $j = 1, 2, \ldots$.

Es seien $u_l(x, y)$ und $u_k(x, y)$ irgend zwei derartige Lösungen. Ihre Werte auf $\mathfrak{K}$ (ihre „Randwerte") wollen wir kurz mit $u_l(s)$ und $u_k(s)$ bezeichnen und die Werte der Normalableitung von u_l bzw. u_k bezüglich $\mathfrak{K}$ mit $v_l(s) = \frac{\partial u_l}{\partial n}$ bzw. $v_k(s) = \frac{\partial u_k}{\partial n}$ $(0 \leqq s \leqq 2\pi)$. Wegen $u_l(s) = 0$, $u_k(s) = 0$ gilt dann die Relation

$$\text{(G)} \qquad \int_0^{2\pi} [u_l(s)\, v_k(s) - u_k(s)\, v_l(s)]\, ds = 0.$$

Wir wollen dafür sagen: „*Die Randwerte* (u_l, v_l) *und* (u_k, v_k) *stehen in G-Relation*".

Allgemein stehen irgend zwei lineare Aggregate von Randwerten: $(\sum_{\mu=1}^{m} c_\mu u_\mu, \sum_{\mu=1}^{m} c_\mu v_\mu)$ bzw. $(\sum_{\nu=1}^{n} \gamma_\nu u_\nu, \sum_{\nu=1}^{n} \gamma_\nu v_\nu)$ mit konstanten Koeffizienten

c_μ, γ_ν in G-Relation; dies drücken wir so aus: „*Die* (u_l, v_l), $l = 1, 2, \ldots$, *bilden ein G-System*". Die (u_l, v_l) nennen wir die „*Elemente*" des Systems.

Unabhängig von diesem Ausgangspunkt legen wir den *Begriff eines G-Systems* fest wie folgt: Gegeben sei eine Menge von Funktionspaaren $(u(s), v(s))$, wo u und v *eindeutige, reellwertige endliche Funktionen der reellen Veränderlichen* s $(0 \leqq s < 2\pi)$ bedeuten, *die nebst ihrem Quadrate im Lebesgueschen Sinne integrierbar* (*summierbar*) *sind*; kurz gesagt: u und v sollen „*Funktionen der Klasse* L^2" sein[1]). Die Menge liefert ein G-System, wenn irgend zwei ihrer „Elemente" (u_1, v_1) und (u_2, v_2) in G-Relation stehen, d. h. wenn

$$\text{(G)} \qquad \int_0^{2\pi} [u_1(s)\, v_2(s) - u_2(s)\, v_1(s)]\, ds = 0.$$

Man beachte dabei, daß das Produkt uv im Lebesgueschen Sinne integrierbar ist, wenn u und v zur Klasse L^2 gehören.

Beispiel eines G-Systems ist die Menge der Elemente (u, v), wobei $u(s) = 0$ in $0 \leqq s < 2\pi$, während $v(s)$ eine *beliebige* Funktion der Klasse L^2 ist. Dieses G-System Γ_0 hat zudem die Eigenschaft, daß es keiner Erweiterung mehr fähig ist im folgenden Sinne: Irgendein zur Klasse L^2 gehöriges Element (u, v), welches zu *allen* Elementen von Γ_0 in G-Relation steht, gehört selbst zu Γ_0, d. h. es ist $u(s) = 0$, $0 \leqq s < 2\pi$. In der Tat: Ist $u = f(s)$, $v = g(s)$ gegeben und soll dieses Element mit allen Elementen von Γ_0 in G-Relation stehen, so muß das insbesondere für das Element $u_1 = 0$, $v_1 = f(s)$ von Γ_0 gelten, also

$$\int_0^{2\pi} [u(s)\, v_1(s) - u_1(s)\, v(s)]\, ds = 0 = \int_0^{2\pi} [f(s)]^2\, ds,$$

woraus $f(s) = 0$ (bis auf eine Nullmenge[1])) in $0 \leqq s < 2\pi$ sich ergibt.

Ein in diesem Sinne vollständiges G-System, wie es Γ_0 ist, nennen wir kurz ein „*komplettes G-System*". Und die, eine „willkürliche" Funktion $\varphi(s)$, d. h. eine beliebige Funktion $\varphi(s)$ der Klasse L^2, enthaltende Darstellung $u(s) = 0$, $v(s) = \varphi(s)$ für die Elemente von Γ_0 heiße die „*Parameterdarstellung von* Γ_0".

---

[1]) Wir betrachten also im folgenden nur *endliche* Funktionen. Dies ist keine wesentliche Einschränkung der Allgemeinheit, da es zu *jeder* über $0 \leqq s < 2\pi$ summierbaren Funktion eine zu ihr *äquivalente*, d. h. höchstens in den Punkten einer Nullmenge von der ersten verschiedene, *endliche* Funktion gibt. Vgl. etwa C. Carathéodory, Vorlesungen über reelle Funktionen (Leipzig 1918, § 387 und 391). Im Interesse einer kurzen Ausdrucksweise werden wir übrigens im folgenden durchweg *zwei* (endliche) *Funktionen als nicht verschieden ansehen, wenn sie äquivalent sind.*

Jedes komplette G-System gibt Anlaß zu einer Greenschen Randbedingung bzw. zu einer Greenschen Randwertaufgabe[2]).

2. Außer Γ_0 sind noch andere Beispiele von kompletten G-Systemen (bzw. Greenschen Randbedingungen) bekannt[2]), die wir kurz anführen wollen:

a) Der Randwertaufgabe, bei welcher $\frac{\partial u}{\partial n} = 0$ gefordert wird, entspricht das komplette G-System mit der Parameterdarstellung $u = \psi(s)$, $v = 0$, unter $\psi(s)$ wieder eine „willkürliche" Funktion verstanden. („Zweite Randwertaufgabe".)

b) Ein komplettes G-System bildet ferner die Gesamtheit aller (zur Klasse L^2 gehörigen) Elemente (u, v), für welche

$$f_1(s)\,u(s) + f_2(s)\,v(s) = 0,$$

unter $f_1(s)$, $f_2(s)$ gegebene beschränkte Funktionen der Klasse L^2 verstanden, für die etwa $f_2(s) \geqq 1$, $(0 \leqq s < 2\pi)$ bleibt. („Sturmsche Aufgabe".) Eine Parameterdarstellung des Systems ist: $u = \varphi(s)$, $v = -\frac{f_1(s)}{f_2(s)}\varphi(s)$.[3]) Man verifiziert wie oben, daß die so dargestellten Elemente tatsächlich ein komplettes G-System bilden.

c) Die „allgemeinste Sturmsche Aufgabe" wird bestimmt durch ein komplettes G-System mit der Parameterdarstellung

$$u(s) = \varphi(s), \qquad v(s) = f(s)\,\varphi(s) + \int_0^{2\pi} K(s,t)\,\varphi(t)\,dt.$$

Dabei ist $f(s)$ eine gegebene beschränkte Funktion der Klasse L^2, so daß auch $f(s)\cdot\varphi(s)$ zur Klasse L^2 gehört[3]); $K(s, t)$ hängt von den beiden reellen Variablen s und t ab, ist symmetrisch in s und t und überdies so beschaffen, daß für willkürliches $\varphi(s)$ der Klasse L^2 auch $\int_0^{2\pi} K(s,t)\,\varphi(t)\,dt$ zur Klasse L^2 gehört; dies ist gewiß der Fall, wenn $\iint [K(s,t)]^2\,ds\,dt$ existiert.

d) Die „pseudoperiodische Randwertaufgabe" entspricht einem G-System, dessen sämtliche Elemente (u, v) den Bedingungen genügen

$$\left.\begin{aligned} u(s+\pi) &= v(s) + \int_0^{\pi} A(s,t)\,u(t)\,dt \\ u(s) &= v(s+\pi) \end{aligned}\right\} \quad \text{für} \quad 0 \leqq s < \pi.$$

---

[2]) Vgl. R. Bär, Über Greensche Randwertaufgaben bei der Schwingungsgleichung (Diss. Würzburg 1915, S. 18—25).

[3]) Da $\frac{f_1(s)}{f_2(s)}$ nach Voraussetzung eine beschränkte, $\varphi(s)$ bzw. $[\varphi(s)]^2$ eine endliche, summierbare Funktion ist, so wird auch $v(s)$ bzw. $[v(s)]^2$ summierbar. Vgl. Carathéodory, l. c. [1]), § 398.

Die gegebene Funktion $A(s,t)$ erfüllt die gleichen Bedingungen wie $K(s,t)$ in c); speziell ist also $A(s,t) = A(t,s)$. Als Parameterdarstellung für die Elemente des Systems hat man

$$\begin{aligned} u(s) &= \varphi(s), && 0 \leqq s < 2\pi, \\ v(s) &= \varphi(s+\pi) - \int_0^\pi A(s,t)\,\varphi(t)\,dt, && 0 \leqq s < \pi, \\ &= \varphi(s-\pi), && \pi \leqq s < 2\pi. \end{aligned}$$

Dabei ist wieder $\varphi(s)$ die „willkürliche" Funktion. Das System erweist sich ebenfalls als komplett.

e) Schließlich sind zu erwähnen die Aufgaben, bei denen längs der Strecke $0 \leqq s < 2\pi$ stückweise verschiedene der eben aufgeführten Randbedingungen vorgeschrieben sind.

3. Allen bisher betrachteten Beispielen ist gemeinsam, daß die betreffenden kompletten G-Systeme eine *Parameterdarstellung* besitzen im folgenden Sinne: Es kann entweder $u(s)$ oder $v(s)$ oder, allgemein zu reden, ein linearer Ausdruck $\alpha(s)\,u(s) + \beta(s)\,v(s)$ als *willkürliche* Funktion von s vorgeschrieben werden: $\varphi(s) = \alpha(s)\,u(s) + \beta(s)\,v(s)$. [4]) Ferner war dann $u(s) = L_1(\varphi)$, $v(s) = L_2(\varphi)$, wobei L_1 und L_2 je eine passend gewählte Schar von Linearoperationen bedeuten. Durch die Parameterdarstellung war das G-System jeweils vollkommen bestimmt.

Um *alle diese Beispiele als spezielle Fälle einer allgemeinen Klasse von kompletten G-Systemen einzuordnen*, haben wir daher folgende *Aufgabe* zu stellen:

Gegeben seien die beiden beschränkten Funktionen $\alpha(s)$ und $\beta(s)$ der Klasse L^2; überdies sei $[\alpha(s)]^2 + [\beta(s)]^2 \geqq 1$, $0 \leqq s < 2\pi$.

Gesucht werden alle (kompletten[5])) G-Systeme (bzw. ihre Parameterdarstellungen), in welchen zu willkürlich vorgegebenem, zur Klasse L^2 gehörigem $\varphi(s)$ ein Element (u, v) existiert, so daß

$$\alpha(s)\,u(s) + \beta(s)\,v(s) = \varphi(s),$$

wobei u und v ebenfalls der Klasse L^2 angehören sollen[6]).

---

[4]) Die über $\alpha(s)$ und $\beta(s)$ zu machenden Annahmen werden weiter unten angegeben.

[5]) Es ergibt sich hinterher, daß das G-System ohne weiteres komplett ist, falls es die übrigen Bedingungen erfüllt (vgl. Nr. 7).

[6]) Zufolge der über $\alpha(s)$ und $\beta(s)$ gemachten Annahmen gehört $\varphi(s)$ zur Klasse L^2, wenn letzteres für $u(s)$ und $v(s)$ zutrifft. Vgl. Carathéodory l. c. [2]).

§ 2.

Lösung.

4. Zur Lösung der Aufgabe sei vorerst bemerkt: Zu $\alpha(s)$ und $\beta(s)$ kann man auf mannigfach verschiedene Arten zwei zur Klasse L^2 gehörige, beschränkte Funktionen $\gamma(s)$ und $\delta(s)$ bestimmen, so daß

$$\alpha(s)\,\delta(s)-\beta(s)\,\gamma(s)=1,\quad 0\leqq s<2\pi.\ ^{7)}$$

Setzt man

$$(\mathrm{S})\qquad \begin{aligned} U(s)&=\alpha(s)\,u(s)+\beta(s)\,v(s),\\ V(s)&=\gamma(s)\,u(s)+\delta(s)\,v(s),\end{aligned}\qquad 0\leqq s<2\pi,$$

so gilt der

Satz. *Einem kompletten G-System (u, v) entspricht vermöge* (S) *wiederum ein komplettes G-System (U, V) und umgekehrt.*

Zum Beweise muß man vorab bemerken, daß zugleich mit $u(s)$ und $v(s)$ auch $U(s)$ und $V(s)$ zur Klasse L^2 gehören[8]).

Ist ferner $U_j=\alpha u_j+\beta v_j$, $V_j=\gamma u_j+\delta v_j$, $j=1, 2$, so gilt

$$\int_0^{2\pi}[U_1V_2-U_2V_1]\,ds=\int_0^{2\pi}(\alpha\delta-\beta\gamma)(u_1v_2-u_2v_1)\,ds=\int_0^{2\pi}(u_1v_2-u_2v_1)\,ds.$$

Es stehen also (U_1, V_1) und (U_2, V_2) in G-Relation, wenn dies für (u_1, v_1) und (u_2, v_2) zutrifft und umgekehrt.

Schließlich ist das G-System (U, V) dann und nur dann komplett, wenn das System der (u, v) komplett ist. Es sei nämlich (U^*, V^*) irgendein zur Klasse L^2 gehöriges Funktionenpaar, welches mit allen Elementen $(U=\alpha u+\beta v,\ V=\gamma u+\delta v)$ in G-Relation steht, unter (u, v) die Elemente eines kompletten G-Systems verstanden. Aus

$$\begin{aligned}&\int_0^{2\pi}[V^*(\alpha u+\beta v)-U^*(\gamma u+\delta v)]\,ds=0\\ =&\int_0^{2\pi}[u(\alpha V^*-\gamma U^*)-v(\delta U^*-\beta V^*)]\,ds\end{aligned}$$

folgt, daß die (zur Klasse L^2 gehörigen) Funktionen $\bar u=\delta U^*-\beta V^*$ und $\bar v=\alpha V^*-\gamma U^*$ ebenfalls dem (kompletten) G-System der (u, v) ange-

---

7) Beispiel: $\gamma(s)=0$, $\delta(s)=\dfrac{1}{\alpha(s)}$ für alle s aus $0\leqq s<2\pi$, für welche $|\alpha(s)|\geqq 1:\sqrt{2}$. $\gamma(s)=-1:\beta(s)$, $\delta(s)=0$ für alle s mit $|\alpha(s)|<1:\sqrt{2}$. Es ist somit $\gamma(s)$ und $\delta(s)$ meßbar auf $0\leqq s<2\pi$. Überdies gilt: $|\delta(s)|\leqq\sqrt{2}$ und $|\gamma(s)|\leqq\sqrt{2}$; letztere Ungleichung folgt aus $|\beta(s)|^2\geqq 1-|\alpha(s)|^2>\frac{1}{2}$ für alle s mit $|\alpha(s)|<1:\sqrt{2}$.

hören; denn sie stehen mit *allen* (u, v) in G-Relation. Demnach gehören $U^* = \alpha\bar{u} + \beta\bar{v}$, $V^* = \gamma\bar{u} + \delta\bar{v}$ zum G-System der (U, V); dieses ist mithin komplett. Ebenso ergibt sich die Umkehrung.

5. Den in Nr. 4 gemachten Feststellungen zufolge haben wir zur Lösung unserer Aufgabe, d. h. zur Bestimmung aller in Nr. 3 charakterisierten, kompletten G-Systeme, lediglich diejenigen Systeme zu betrachten, für welche $\alpha(s) \equiv 1$, $\beta(s) \equiv 0$ ist. Denn das (vermöge (S) transformierte) System (U, V) ist so beschaffen, daß $U(s) = \alpha(s)\,u(s) + \beta(s)\,v(s) = \varphi(s)$ willkürlich (aus der Klasse L^2) gewählt werden kann, und wenn alle derartigen G-Systeme bekannt sind, erhält man durch eine Transformation (S) das allgemeine G-System von der in Nr. 3 festgelegten Art.

Zur Lösung der gestellten Aufgabe müssen wir also lediglich *feststellen, auf welche Weise in einem kompletten*[5]) *G-System von Elementen (U, V), in welchem $U(s)$ willkürlich (aus der Klasse L^2) vorgeschrieben werden kann, die Funktion $V(s)$ von $U(s)$ abhängt.*

6. Es sei ein G-System (U, V) der eben erwähnten Beschaffenheit gegeben. Wir wählen für die willkürliche Funktion $U(s)$ der Reihe nach die (sicher zur Klasse L^2 gehörigen) Elemente $f_1(s)$, $f_2(s)$, ... eines normierten, vollständigen Orthogonalsystems von stetigen Funktionen in bezug auf das Intervall $0 \leqq s \leqq 2\pi$. Es seien bezeichnet: Eine, zu $U(s) = f_\varrho(s)$ im gegebenen G-System gehörige, Funktion $V(s)$ mit $g_\varrho(s)$, ferner die Fourierkoeffizienten von $U(s)$ bzw. $V(s)$ bzw. $g_\nu(s)$ bezüglich des Orthogonalsystems der $f_\varrho(s)$ mit U_ϱ bzw. V_ϱ bzw. $g_{\nu\varrho}$ $(\nu, \varrho = 1, 2, \ldots)$.

Die G-Relation zwischen $(f_\varrho(s), g_\varrho(s))$ und einem beliebigen anderen Elemente $(U(s), V(s))$ des gegebenen G-Systems liefert

$$0 = \int_0^{2\pi} [f_\varrho(s)V(s) - g_\varrho(s)U(s)]\,ds = \int_0^{2\pi} f_\varrho(s)V(s)\,ds - \int_0^{2\pi} g_\varrho(s)U(s)\,ds.$$

Bedenkt man, daß das erste Integral rechter Hand V_ϱ ist, und wendet auf das zweite Integral den Parsevalschen Satz an, so kommt:

$$(\mathrm{G}_\varrho^*) \qquad V_\varrho = \sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu \qquad (\varrho = 1, 2, \ldots).$$

Da sämtliche vorkommende Funktionen zur Klasse L^2 gehören, konvergieren $\sum_{\nu=1}^{\infty} U_\nu^2$, $\sum_{\nu=1}^{\infty} g_{\varrho\nu}^2$; und daher konvergiert auch $\sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu$. Umgekehrt gehört nach dem Riesz-Fischerschen Satze zu jeder Folge reeller Zahlen $\tilde{U}_1, \tilde{U}_2, \ldots$ mit konvergenter Quadratsumme eine Funktion $\tilde{U}(s)$ der

Klasse L^2, für die $\int\limits_0^{2\pi} \tilde{U}(s) f_\varrho(s)\,ds = \tilde{U}_\varrho$.[8]) Daraus folgt, daß die in (G_ϱ^*) auftretenden Koeffizienten U_ν nur konvergente Quadratsumme zu besitzen brauchen und im übrigen beliebig sind, entsprechend der Willkürlichkeit von $U(s)$. Weiter folgt, daß die zu *jedem* beliebigen derartigen System von Zahlen $U_1, U_2, \ldots$ aus (G_ϱ^*) sich ergebenden Zahlen V_ϱ, als Fourierkoeffizienten einer Funktion der Klasse L^2, wiederum konvergente Quadratsumme besitzen müssen. Die Bilinearform, welche durch das System der $g_{\nu\varrho}$ $(\nu, \varrho = 1, 2, \ldots)$ definiert wird, ist mithin *beschränkt*[9]). Die eben genannte Bilinearform $\sum\limits_{\varrho\nu} g_{\varrho\nu} x_\varrho y_\nu$ muß überdies *symmetrisch* sein, d. h. es muß $g_{\varrho\nu} = g_{\nu\varrho}$ gelten $(\nu, \varrho = 1, 2, \ldots)$, falls die Gleichungen (G_ϱ^*) zu einem G-System gehören sollen. Betrachtet man nämlich zwei Elemente $(U(s), V(s))$ und $(U^*(s), V^*(s))$ des vorgelegten G-Systems, so liefert die G-Relation zwischen (U, V) und (U^*, V^*):

$$0 = \int\limits_0^{2\pi} (UV^* - U^*V)\,ds = \sum_{\varrho=1}^{\infty} (U_\varrho V_\varrho^* - U_\varrho^* V_\varrho).$$

Setzt man für V_ϱ und V_ϱ^* ihre Werte aus (G_ϱ^*) ein, so ergibt sich:

$$(G^{**}) \qquad \begin{aligned} 0 &= \sum_{\varrho=1}^{\infty} \Big[U_\varrho \Big(\sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu^*\Big) - U_\varrho^* \Big(\sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu\Big)\Big] \\ &= \sum_{\varrho=1}^{\infty} U_\varrho \Big[\sum_{\nu=1}^{\infty} (g_{\varrho\nu} - g_{\nu\varrho}) U_\nu^*\Big].\ ^{10)} \end{aligned}$$

Wählt man jetzt speziell $U(s)$ und $U^*(s)$ so, daß ihre Fourierkoeffizienten mit Ausnahme je eines einzigen sämtlich verschwinden, setzt also etwa:

$$\begin{aligned} U_\varrho &= 0, \ \varrho \neq k_1, & U_\varrho^* &= 0, \ \varrho \neq k_2, \\ &= 1, \ \varrho = k_1, & &= 1, \ \varrho = k_2, \end{aligned}$$

wo k_1 und k_2 voneinander verschiedene natürliche Zahlen bedeuten, so liefert die letzte Beziehung:

$$g_{k_1 k_2} - g_{k_2 k_1} = 0.$$

Umgekehrt läßt (G^{**}) erkennen, daß die Gleichungen (G_ϱ^*) immer dann zu einem G-System Anlaß geben, falls die Bilinearform $\sum\limits_{\nu\varrho} g_{\nu\varrho} x_\nu y_\varrho$ beschränkt und symmetrisch ist.

---

[8]) $\tilde{U}(s)$ kann immer als endlich vorausgesetzt werden. Vgl. Fußnote [1]).

[9]) E. Hellinger und O. Toeplitz, Grundlagen für eine Theorie der unendlichen Matrizen. Math. Ann. 69 (1910), § 10. Vgl. auch F. Riesz, Les systèmes d'équations linéaires à une infinité d'inconnues (Paris 1913, S. 81).

[10]) Die Summationen nach ν und ϱ sind miteinander vertauschbar, weil die Bilinearform beschränkt ist. Vgl. Hellinger-Toeplitz, l. c. [9]), § 2.

Jedes durch Vermittelung einer derartigen Bilinearform definierte G-System ist von selbst komplett. Sei nämlich irgendein der Klasse L^2 angehöriges Element $(U^*(s), V^*(s))$ gegeben, welches zu allen Elementen eines G-Systems der eben erwähnten Art in G-Relation steht, so liefert die G-Relation bei Benutzung der Beziehungen (G_ϱ^*) wie oben:

$$0 = \sum_{\varrho=1}^{\infty}[U_\varrho V_\varrho^* - V_\varrho U_\varrho^*] = \sum_{\varrho=1}^{\infty}[U_\varrho V_\varrho^* - U_\varrho^*(\sum_{\varrho=1}^{\infty} g_{\varrho\nu} U_\nu)]$$

$$= \sum_{\varrho=1}^{\infty} U_\varrho [V_\varrho^* - \sum_{\varrho=1}^{\infty} g_{\varrho\nu} U_\nu^*].$$

Wählt man wieder $U(s)$ so, daß etwa: $U_\varrho = 0$, $\varrho \neq k$; $U_k = 1$, so folgt

$$V_k^* = \sum_{\nu=1}^{\infty} g_{k\nu} U_\nu^* \qquad (k = 1, 2, \ldots),$$

und $(U^*(s), V^*(s))$ ist demnach selbst ein Element unseres G-Systems, w. z. b. w.

Zusammenfassend können wir somit sagen: *Ein System von Elementen* $(U(s), V(s))$, *wo* $U(s)$ *irgendeine willkürliche Funktion* (*der Klasse* L^2) *sein kann, ist dann und nur dann ein komplettes G-System, wenn die Fourierkoeffizienten von* $V(s)$ *durch lineare Transformation aus den Fourierkoeffizienten von* $U(s)$ *hervorgehen und wenn dabei diejenige Bilinearform* (*in unendlich vielen Variablen*), *welche mit den Koeffizienten jener Transformation gebildet wird, beschränkt und symmetrisch ist.*

7. Durch *Rückgang zu den Funktionen*[11]) erhält man jetzt die *Parameterdarstellung* für unsere kompletten G-Systeme.

Wir gehen aus von den Beziehungen

$$(G_\varrho^*) \qquad V_\varrho = \sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu \qquad (\varrho = 1, 2, \ldots)$$

und bilden

$$(1) \qquad \sum_{\varrho=1}^{\infty} V_\varrho l_\varrho(s) = \sum_{\varrho=1}^{\infty}(\sum_{\nu=1}^{\infty} g_{\varrho\nu} U_\nu) l_\varrho(s) = \sum_{\nu=1}^{\infty} U_\nu [\sum_{\nu=1}^{\infty} g_{\nu\varrho} l_\varrho(s)];$$

dabei ist gesetzt

$$(\mathrm{I}) \qquad l_\varrho(s) = \int_0^s f_\varrho(\sigma)\, d\sigma \qquad (\varrho = 1, 2, \ldots);$$

und wegen $\sum_{\varrho=1}^{\infty}(l_\varrho(s))^2 \leqq 1$ konvergiert $\sum_{\varrho=1}^{\infty} V_\varrho l_\varrho(s)$, während die weiter vor-

[11]) Nach Hellinger, Die Orthogonalinvarianten quadratischer Formen von unendlich vielen Variablen. Diss. Göttingen 1907, Anhang (§ 12), S. 83/84.

genommene Vertauschung der Summationsfolgen wegen der Beschränktheit der Bilinearform $\sum_{\nu\varrho} g_{\nu\varrho} x_\nu y_\varrho = B(x, y)$ zulässig ist[10]).

Nun gibt es aber eine zur Klasse L^2 gehörige Funktion $\Phi(s, t)$ von t, welche außerdem von s abhängt und die Eigenschaft hat, daß[12])

$$(2) \qquad \int_0^{2\pi} \Phi(s, t)\, U(t)\, dt = \sum_{\nu=1}^{\infty} \Big[\sum_{\varrho=1}^{\infty} g_{\nu\varrho} l_\varrho(s)\Big] U_\nu .$$

Bei diesem Schlusse ist ebenfalls die Beschränktheit von $B(x, y)$ benutzt. Des näheren ist (fast überall):[11])

$$(\mathrm{II}) \qquad \Phi(s, t) = \frac{d}{dt}[K(s, t)], \quad \text{wobei} \quad K(s, t) = \sum_{\nu=1}^{\infty} \sum_{\varrho=1}^{\infty} g_{\nu\varrho} l_\nu(s)\, l_\varrho(t).$$

Aus (1) und (2) folgt

$$(3) \qquad \sum_{\varrho=1}^{\infty} V_\varrho l_\varrho(s) = \int_0^{2\pi} \Phi(s, t)\, U(t)\, dt .$$

Aus (3) ergibt sich schließlich[13]) (höchstens mit Ausnahme einer Nullmenge):

$$(\mathrm{III}) \qquad V(s) = \frac{d}{ds}\Big[\int_0^{2\pi} \Phi(s, t)\, \varphi(t)\, dt\Big], \quad U(s) = \varphi(s) \qquad (0 \leqq s < 2\pi).$$

Und dies ist die gewünschte Parameterdarstellung.

Zusammenfassung. *Jedes G-System der in Nr.* 5 *festgelegten Art ist charakterisiert durch eine Funktion $K(s, t)$ von der in Formel* (II) *angegebenen Form, wobei die Konstanten $g_{\nu\varrho}$ die Koeffizienten einer beschränkten symmetrischen Bilinearform darstellen und die Funktionen $l_\nu(s)$ durch* (I) *definiert sind. Vermittelst $K(s, t)$ ergibt sich die Parameterdarstellung des Systems in der durch* (II) *und* (III) *bezeichneten Weise. Das G-System ist komplett.*

---

[12]) Hahn, Über die Integrale des Herrn Hellinger usw. Monatshefte f. Math. u. Physik 23 (1912), S. 184/185.

[13]) Hellinger, l. c. [11]), S. 81/82.

(Eingegangen am 26. 11. 1923.)

Ein Satz über Dirichletsche Reihen.

Von

Werner Rogosinski in Königsberg i. Pr.

Im folgenden bezeichne die formal gebildete

$$(1)\qquad \sum_{n=0}^{\infty} a_n e^{-\lambda_n s}$$

mit

$$(2)\qquad 0 \leqq \lambda_0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n < \lambda_{n+1} < \ldots; \qquad \lambda_n \to \infty$$

und komplexen Koeffizienten $a_n \neq 0$ eine allgemeine Dirichletsche Reihe der komplexen Variablen

$$(3)\qquad s = \sigma + i t.$$

Wenn für irgendein $\varkappa \geqq 0$ und ein festes s der Grenzwert

$$(4)\qquad \lim_{x \to \infty} \sum_{\lambda_n \leqq x} a_n e^{-\lambda_n s} \left(1 - \frac{\lambda_n}{x}\right)^{\varkappa} = f(s)$$

existiert, so soll für dieses s die Reihe (1) nach Riesz[1]) $(\lambda, \varkappa)$-summierbar zum Werte $f(s)$ heißen. $(\lambda, 0)$-Summabilität bedeutet also Konvergenz von (1).

Herr Bohr[2]) hat 1913 den folgenden, bei den damaligen Kenntnissen über die Reihen (1) sehr merkwürdigen Satz bewiesen:

Satz 1. *Es sei* (1) *in einer Halbebene $\sigma > \alpha$ absolut konvergent. Das durch* (1) *hier erklärte Funktionselement $f(s)$ sei in der Viertelebene*

$$(5)\qquad \sigma > \gamma; \qquad t > \tau$$

---

[1]) Vgl. etwa G. H. Hardy und M. Riesz, The General Theory of Dirichlet's Series, Cambridge Tracts 18 (1915), S. 21 (*hinfort „H. R." zitiert*).

[2]) H. Bohr, Ein Satz über Dirichletsche Reihen, Sitzungsberichte der Königl. Bayerischen Akademie der Wissenschaften, Mathem. physikalische Klasse (1913), S. 557–562.

regulär, und daselbst für ein $M > 0$

(6) $$|f(s)| \leqq M$$

erfüllt.

Dann ist $f(s)$ *sogar in der ganzen Halbebene* $\sigma > \gamma$ *regulär, und daselbst gilt* (6).

Neuere Untersuchungen des Herrn Bohr[3]) über die „*Quasiperiodizität*" der Reihen (1) machen die Aussage dieses Satzes im Rahmen der Theorie viel einleuchtender und legen seine Verallgemeinerungsfähigkeit nahe. In der Tat läßt er sich, und zwar in der Hauptsache unter Beibehaltung der Bohrschen Beweismethode, die auf der Anwendung von *diophantischen Approximationen* beruht, sehr wesentlich erweitern. Insbesondere erweist sich die Voraussetzung einer Halbebene *absoluter Konvergenz* als überflüssig. Ich behaupte nämlich den folgenden

Satz 2. *Es sei* (1) *in einer Halbebene* $\sigma > \alpha$ *für irgendein* $\varkappa \geqq 0$ $(\lambda, \varkappa)$*-summierbar. Das durch* (4) *hier erklärte Funktionselement* $f(s)$ *sei in der Viertelebene* (5) *regulär, und daselbst für zwei komplexe Werte* $a \neq b$

(7) $$f(s) \neq a \quad \text{und} \quad \neq b.$$

Dann ist $f(s)$ *sogar in der ganzen Halbebene* $\sigma > \gamma$ *regulär und daselbst* (7) *erfüllt.*

Zunächst möchte ich an einige, im folgenden benötigte, einfache Haupttatsachen der durch (4) erklärten Summabilitätsmethode erinnern.

1. Wenn (1) in einem Punkte $s = s_0 = \sigma_0 + i\,t_0$ für ein $\varkappa \geqq 0$ $(\lambda, \varkappa)$-summierbar ist, so auch in der ganzen Halbebene $\sigma > \sigma_0$, und zwar gleichmäßig in jedem Winkelraum

$$|\,\text{ampl.}\,(s - s_0)\,| \leqq \vartheta < \frac{\pi}{2}.$$

Das durch (4) in $\sigma > \sigma_0$ erklärte Funktionselement $f(s)$ ist also daselbst regulär[4]).

2. Wenn (1) in s_0 $(\lambda, \varkappa)$-summierbar ist zum Werte $f(s_0)$, so auch $(\lambda, \varkappa_1)$ mit $\varkappa_1 > \varkappa$ zum gleichen Werte[5]).

3. Wenn (1) in s_0 $(\lambda, \varkappa)$-summierbar ist, so gilt für jedes feste $\varepsilon > 0$

(8) $$f(s) = o(|t|^{\varkappa+1})$$

gleichmäßig in $\sigma \geqq \sigma_0 + \varepsilon$.[6])

[3]) H. Bohr, Über eine quasi-periodische Eigenschaft Dirichletscher Reihen ..., Math. Annalen 85 (1922), S. 115—122.

[4]) H. R., Theorem 23, S. 39.

[5]) H. R., Theorem 16, S. 29.

[6]) H. R., Theorem 38, S. 49.

Nunmehr will ich zuerst zeigen, daß unsere Voraussetzung (7) nicht wesentlich mehr als (6) besagt. Es gilt nämlich in leichter Modifikation eines Satzes von Herrn Wennberg[7]) der folgende

Hilfssatz. *Es sei* δ *und* $\varepsilon > 0$. *Unter den Voraussetzungen von Satz 2 gibt es ein* $M = M(\delta, \varepsilon) > 0$, *so daß in jeder Viertelebene*

$$\sigma \geqq \gamma + \varepsilon; \qquad t \geqq \tau + \delta$$

sogar (6) *erfüllt ist.*

Beweis. Ohne Beschränkung der Allgemeinheit sei

$$\text{(a)} \qquad a = 0 \quad \text{und} \quad b = 1,$$

anderenfalls man $\frac{f(s) - a}{b - a}$ statt $f(s)$ betrachte.

Für $\sigma > \alpha$ ist

$$\text{(b)} \qquad f(s) = a_0 e^{-\lambda_0 s}\left\{1 + e^{-(\lambda_1 - \lambda_0)s} \cdot \lim_{x \to \infty} \sum_{\lambda_1 \leqq \lambda_n \leqq x} \frac{a_n}{a_0}\left(1 - \frac{\lambda_n}{x}\right)^{\varkappa} e^{-(\lambda_n - \lambda_1)s}\right\}$$
$$= a_0 e^{-\lambda_0 s}\{1 + e^{-(\lambda_1 - \lambda_0)s} \cdot \varphi(s)\}.$$

Hierin ist, da $f(s)$ nach Bemerkung 1 jedenfalls auf der Halbgeraden

$$\text{(c)} \qquad \sigma \geqq \alpha + 1; \qquad t = \tau + \delta$$

gleichmäßig $(\lambda, \varkappa)$-summabel ist, daselbst für ein passendes $K > 0$

$$\text{(d)} \qquad |\varphi(s)| < K.$$

Es bezeichne nunmehr

$$\text{(e)} \qquad \zeta = J(z) = \frac{g_2^3}{g_2^3 - 27\, g_3^2}$$

mit

$$\text{(f)} \qquad g_2 = 60 \sum_{\substack{m,n \\ |m|+|n| \neq 0}}{}' \frac{1}{(m + n z)^4}; \qquad g_3 = 140 \sum_{\substack{m,n \\ |m|+|n| \neq 0}}{}' \frac{1}{(m + n z)^6}$$

die „Modulfunktion“[8]), welche die Halbebene $\mathfrak{J}(z) > 0$, wo sie regulär ist, auf die bei $\zeta = 0, 1, \infty$ verzweigte, unendlichblättrige „Modulfläche“ der ζ-Ebene abbildet.

Die inverse Funktion $z = \omega(\zeta)$ ist umgekehrt für alle $\zeta \neq 0, 1, \infty$ regulär, jedoch unendlich-vieldeutig. Stets ist aber $\mathfrak{J}(\omega) > 0$.

---

[7]) Sven Wennberg, Zur Theorie der Dirichletschen Reihen, Inauguraldissertation, Uppsala (1920). Vgl. Satz VI auf S. 30, dessen Beweis hier, nur für unsere Zwecke zurecht gemacht, wiedergegeben ist.

[8]) Vgl. etwa Hurwitz-Courant, Allgemeine Funktionentheorie und elliptische Funktionen, Berlin bei Springer (1922), II. Abschn., 4. Kapitel, S. 206–216 und III. Abschn., 4. Kapitel, § 9, S. 312–317.

Nun kann man bekanntlich[8]) einen solchen „Zweig" von $\omega(\zeta)$ herausgreifen, daß

$$\text{(g)}\quad \omega(\zeta)-\omega(\zeta_0) = (\zeta-\zeta_0)\mathfrak{P}_1(\zeta-\zeta_0)\ \text{für}\ \begin{array}{l}\zeta_0 \neq 0,1\\ |\zeta-\zeta_0| < \mathrm{Min}\{|\zeta_0|,|\zeta_0-1|\},\end{array}$$

$$\text{(h)}\quad \omega(\zeta)-\frac{-1+i\sqrt{3}}{2} = \zeta^{\frac{1}{3}}\mathfrak{P}_2(\zeta) \qquad \text{für}\ |\zeta|<1;\ \zeta \neq 0,$$

$$\text{(i)}\quad \omega(\zeta)-i = \sqrt{\zeta-1}\,\mathfrak{P}_3(\zeta-1)\ \text{für}\ |\zeta-1|<1;\ \zeta\neq 1$$

wird, wo $\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3$ gewisse Potenzreihen ihres Arguments symbolisieren.

Die Funktion

$$\text{(j)}\qquad g(s)=\omega(f(s))$$

ist in der Viertelebene (5) unserer Voraussetzung nach regulär, aber natürlich unendlich vieldeutig. Ein willkürlich herausgegriffener Zweig von ω führt aber zu einer in (5) eindeutigen regulären Funktion. Stets ist in (5)

$$\text{(k)}\qquad \mathfrak{I}\, g(s) > 0.$$

I. $\lambda_0 = 0$. Auf der Halbgeraden (c) ist nach (b) und (d)

$$\text{(l)}\qquad |f(s)-a_0| < |a_0|\cdot K\cdot e^{-\lambda_1\sigma}.$$

Es gibt daher ein $\beta_1 \geqq \alpha+1$, so daß insbesondere für $\sigma \geqq \beta_1$, $t=\tau+\delta$

$$\text{(m)}\qquad |f(s)-a_0| < \begin{cases}\frac{1}{2}\,\mathrm{Min}\{|a_0|,|a_0-1|\}, & \text{falls}\quad a_0\neq 1,\\ \frac{1}{2}, & \phantom{\text{falls}}\quad a_0=1.\end{cases}$$

Für diese s ist dann nach (g), (i) und (l), nachdem der Zweig von ω fixiert ist, für ein gewisses $K_1 > 0$

$$\text{(n)}\qquad \begin{array}{lll} |g(s)-\omega(a_0)| < e^{-\lambda_1\sigma}\cdot K_1, & & a_0 \neq 1,\\ & \text{falls} & \\ |g(s)-i| \quad < e^{-\lambda_1\frac{\sigma}{2}}K_1, & & a_0 = 1.\end{array}$$

II. $\lambda_0 > 0$. Jetzt gilt auf (c)

$$\text{(o)}\qquad |f(s)| < |a_0|\cdot e^{-\lambda_0\sigma}\{1+e^{-(\lambda_1-\lambda_0)\sigma}\cdot K\}.$$

Für $\sigma \geqq \beta_2$, $t=\tau+\delta$ wird daher

$$\text{(p)}\qquad |f(s)| < \frac{1}{2}$$

und also nach (h) für den betrachteten Zweig von ω und ein gewisses $K_2 > 0$

$$\text{(q)}\qquad \left|g(s)-\frac{-1+i\sqrt{3}}{2}\right| < e^{-\lambda_0\frac{\sigma}{3}}\cdot K_2.$$

Es bezeichne nunmehr, je nach dem vorliegenden Falle, $h(s)$ die in (n) bzw. (q) links auftretende Funktion, $\mu > 0$ den daselbst rechts auf-

tretenden Faktor von $-\sigma$ und K_3 die Konstante K_1 oder K_2, so daß also für alle $\sigma \geqq \beta_3$, $t = \tau + \delta$ bei genügend großem β_3

(r) $$|h(s)| < e^{-\mu\sigma} K_3$$

gilt.

Die Funktion

(s) $$F(s) = e^{\mu s}(e^{ih(s)} - 1)$$

ist in der Viertelebene (5) regulär, und daselbst gilt unter Beachtung von (k) die Abschätzung

(t) $$F(s) = O(e^{\mu\sigma}).$$

Insbesondere ist also $F(s)$ auf der vertikalen Halbgeraden $\sigma = \beta_3$, $t \geqq \tau + \delta$ gleichmäßig beschränkt. Dasselbe ist aber auch auf der horizontalen Halbgeraden $\sigma \geqq \beta_3$, $t = \tau + \delta$ der Fall, wie aus (r) unmittelbar folgt. Ein bekanntes Theorem von Lindelöf-Phragmén[9]) lehrt jetzt unter Beachtung von (t), daß $F(s)$ sogar in der ganzen Viertelebene $\sigma \geqq \beta_3$, $t \geqq \tau + \delta$ gleichmäßig beschränkt ist. Es gibt also ein $K_4 > 0$, so daß daselbst

(u) $$|e^{ih(s)} - 1| \leqq K_4 e^{-\mu\sigma}$$

gilt. Da nun nach (r) auf $\sigma \geqq \beta_3$, $t = \tau + \delta$ mit wachsendem σ $h(s) \to 0$ gilt, so folgt aus (u), daß $h(s)$ sogar in unserer Viertelebene mit wachsendem σ gleichmäßig gegen 0 strebt. $g(s)$ muß also daselbst, je nach dem vorliegenden Falle, gleichmäßig gegen $\omega(a_0)$, i oder $\frac{-1+i\sqrt{3}}{2}$ mit wachsendem σ streben. Daraus folgt nun aber weiter, daß

(v) $$f(s) = J\big(\omega(f(s))\big) = J(g(s))$$

gleichmäßig daselbst gegen a_0 strebt; jedenfalls gibt es also ein $M_1 = M_1(\delta) > 0$, so daß bei genügend großem β_4 in der Viertelebene $\sigma \geqq \beta_4$, $t \geqq \tau + \delta$

(w) $$|f(s)| \leqq M_1$$

gilt.

---

[9]) E. Lindelöf et E. Phragmén, Sur une extension d'un principe classique de l'analyse ..., Acta Math. **31** (1908); vgl. S. 385—386. Wir können den Satz so formulieren:

„Es sei $F(s)$ im Innern und auf dem Rande eines Winkelraums der Öffnung ψ in der s-Ebene eindeutig regulär. Auf den Schenkeln gelte für ein $K > 0$, daß $|F(s)| \leqq K$ ist, während im Innern mit wachsendem r für jedes $\varepsilon > 0$ eine gleichmäßige Abschätzung $F(s) = O\left(e^{\varepsilon r^{\frac{\pi}{\psi}}}\right)$ gilt, wo r den Abstand von s gegen den Scheitel bezeichnet. Dann ist sogar im ganzen Winkelraum $|F(s)| \leqq K$ richtig."

In unserem Falle ist $\psi = \frac{\pi}{2}$, und die Abschätzung (t) kann für jedes $\varepsilon > 0$ in der Form $F(s) = O(e^{\varepsilon r^2})$ geschrieben werden.

Wir dürfen uns weiterhin auf $\beta_4 > \gamma + \varepsilon$ beschränken. Dann ist die Existenz eines $M_2 = M_2(\varepsilon, \delta) > 0$ evident, so daß in dem Rechteck

$\gamma + \varepsilon \leqq \sigma \leqq \beta_4$, $\tau + \delta \leqq t \leqq \tau + \delta + \left(\beta_4 - \gamma - \frac{\varepsilon}{2}\right)$

$$(\mathrm{x}) \qquad |f(s)| \leqq M_2$$

richtig ist.

Jeder Punkt des restlichen Halbstreifens

$$(\mathrm{y}) \qquad \gamma + \varepsilon \leqq \sigma \leqq \beta_4; \qquad \tau + \delta + \left(\beta_4 - \gamma - \frac{\varepsilon}{2}\right) \leqq t$$

liegt nun im Innern oder auf dem Rande eines Kreises $\mathfrak{K}(t)$, der um den Punkt $\beta_4 + it$ mit dem Radius $\beta_4 - (\gamma + \varepsilon)$ geschlagen ist. In dem konzentrischen größeren Kreise mit dem Radius $\beta_4 - \left(\gamma + \frac{\varepsilon}{2}\right)$ ist $f(s)$ regulär, und daselbst ist $f(s) \neq 0$ und $\neq 1$. Im übrigen beachte man (w).

Aus einem Zusatz von Herrn Landau[10]) zu dem bekannten Schottkyschen Satze ergibt sich nun unmittelbar die Existenz eines $M_3 = M_3(\varepsilon, \delta) > 0$, so daß in (y)

$$(\mathrm{z}) \qquad |f(s)| \leqq M_3$$

gilt.

Mit (w), (x) und (z) ist aber die Behauptung bewiesen.

Wir beginnen nunmehr mit dem

Beweis von Satz 2. Durch eine einfache Variablentransformation können wir offenbar die Normierung

$$(\mathrm{a}) \qquad \alpha = \tau = 0$$

verwirklichen. Mit $\beta = \mathrm{Max}(0, \gamma)$ ist dann (1) für $\sigma > \beta$ und ein gewisses $\varkappa \geqq 0$ nach Voraussetzung $(\lambda, \varkappa)$-summierbar. Daher gilt gleichmäßig in $\sigma \geqq \beta + \frac{1}{4}$ nach (8) die Abschätzung

$$(\mathrm{b}) \qquad f(s) = o(|t|^{\varkappa+1}).$$

Ferner gibt es nach unserem Hilfssatze ein $M(\varepsilon) > 0$, so daß in der Viertelebene $\sigma \geqq \gamma + \varepsilon$, $t \geqq 1$, wo $0 < \varepsilon \leqq \frac{1}{4}$ sei,

$$(\mathrm{c}) \qquad |f(s)| \leqq M$$

gilt.

---

[10]) C. Carathéodory und E. Landau, Beiträge zur Konvergenz von Funktionenfolgen, Sitzungsberichte der Berliner Akademie (1911), Satz V auf S. 596. Wir wollen ihn so formulieren:

„*Es sei für* $|z| < 1$ *die Funktion* $f(z)$ *eindeutig regulär,* $\neq 0$ *und* $\neq 1$, $|f(0)| \leqq C$. *Dann gibt es ein nicht von* $f(z)$ *abhängendes* $M = M(\vartheta, C) > 0$, *so daß für* $|z| \leqq \vartheta$ *mit* $0 \leqq \vartheta < 1$ *die Beschränkung* $|f(z)| \leqq M$ *gilt.*“

Nun können wir für $\sigma \geqq \beta + \frac{1}{2}$, alle reellen x und ganzes $m \geqq \varkappa + 2$ die Identität[11])

$$\text{(d)} \qquad \sum_{\lambda_n \leqq x} a_n e^{-\lambda_n s}(1 - e^{\lambda_n - x})^m = \frac{m!}{2\pi i}\int\limits_{(1)} \frac{e^{x\zeta} f(\zeta + s)}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)}\, d\zeta$$

ansetzen, in der das Integral rechts, über die Gerade $\Re(\zeta) = 1$ erstreckt, wegen (b) gewiß sogar absolut konvergiert.

Nach Cauchy können wir hierin wegen (b) das Integral unter Beachtung des Residuums des Integranden für $\zeta = 0$ auf die Gerade $\Re(\zeta) = -\frac{1}{4}$ übertragen, und erhalten

$$\text{(e)} \qquad \sum_{\lambda_n \leqq x} a_n e^{-\lambda_n s}(1 - e^{\lambda_n - x})^m = f(s) + \frac{m!}{2\pi i}\int\limits_{(-\frac{1}{4})} \frac{e^{x\zeta} f(\zeta + s)}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)}\, d\zeta.$$

Es ist aber wegen (b) gleichmäßig für alle $\sigma \geqq \beta + \frac{1}{2}$ und alle x bei festem m

$$\text{(f)} \qquad \int\limits_{(-\frac{1}{4})} \frac{e^{x\zeta} f(\zeta + s)}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)}\, d\zeta$$

$$= O\left(e^{-\frac{x}{4}} \int\limits_{-\infty}^{\infty} \frac{|\tau + t|^{\varkappa+1}}{\left|-\frac{1}{4}+i\tau\right|\left|\frac{3}{4}+i\tau\right|\left|\frac{7}{4}+i\tau\right|\ldots\left|\frac{4m-1}{4}+i\tau\right|}\, d\tau\right)$$

$$= O\left(e^{-\frac{x}{4}} |t|^{\varkappa+1}\right),$$

und in der Viertelebene $\sigma \geqq \beta + \frac{1}{2}$, $t \geqq 1$ sogar gleichmäßig wegen (c)

$$\text{(g)} \qquad \int\limits_{(-\frac{1}{4})} \frac{e^{x\zeta} f(\zeta + s)}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)}\, d\zeta$$

$$= O\left(e^{-\frac{x}{4}}\left\{\int\limits_{-\infty}^{-t} \frac{|\tau|^{\varkappa+1}}{\left|-\frac{1}{4}+i\tau\right|\left|\frac{3}{4}+i\tau\right|\ldots\left|\frac{4m-1}{4}+i\tau\right|}\, d\tau\right.\right.$$

$$\left.\left. + \int\limits_{-t}^{\infty} \frac{d\tau}{\left|-\frac{1}{4}+i\tau\right|\left|\frac{3}{4}+i\tau\right|\ldots\left|\frac{4m-1}{4}+i\tau\right|}\right\}\right)$$

$$= O\left(e^{-\frac{x}{4}}\right).$$ [12])

[11]) H. R., Theorem 40 auf S. 51. Wir arbeiten hier also mit Rieszschen (l, m)-Mitteln, was wegen Theorem 30 auf S. 45 natürlich gestattet ist. Die Identität (d) gilt sogar für $m \geqq \varkappa$. Beachte auch unsere Bemerkung 2.

[12]) Dies ist die entscheidende Stelle des Beweises, von der an der Bohrsche Beweisgang kaum verändert einsetzt.

Es sei nunmehr $s_1 = \sigma_1 + i t_1$ ein beliebiger Punkt mit $\sigma_1 > \gamma$. Unser Satz ist bewiesen, wenn wir zeigen können, daß $f(s)$ in s_1 regulär und daselbst $\neq a$ und $\neq b$ ist.

Es bezeichne $\sigma_1' = \mathrm{Max}(\sigma_1, \beta + 1)$, $s_1' = \sigma_1' + i t_1$, $\mathfrak{K}$ den Kreis um s_1' mit dem Radius $\varrho = \sigma_1' - \gamma - \varepsilon$, wo das bei (c) auftretende ε noch überdies $< \sigma_1 - \gamma$ gewählt werde, $\mathfrak{K}'$ den konzentrischen kleineren Kreis mit dem Radius $\frac{1}{2}$. s_1 liegt jedenfalls im Innern von $\mathfrak{K}$.

Ferner sei mit

(h) $$\eta_1, \eta_2, \ldots, \eta_k, \ldots \to 0$$

eine positive Nullfolge vorgelegt.

Für zwei Punkte $s = \sigma + it$ und $s' = \sigma + it'$ mit $\sigma \geqq \sigma_1' - \frac{1}{2}$ gilt nach (e)

(i) $$\sum_{\lambda_n \leqq x} a_n e^{-\lambda_n \sigma} (e^{-\lambda_n i t} - e^{-\lambda_n i t'}) (1 - e^{\lambda_n - x})^m$$
$$= f(s) - f(s') + \frac{m!}{2\pi i} \int\limits_{(-\frac{1}{4})} \frac{e^{x\zeta} \{f(\zeta + s) - f(\zeta + s')\}}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)} d\zeta.$$

Ist hierin s' auf den Kreis $\mathfrak{K}'$ und s auf $t \geqq 1$ beschränkt, so folgt aus (f) und (g), da ja bei festem s_1 die $|t'|$ beschränkt sind, gleichmäßig in diesen s und s' für ein geeignetes $x = x_k = x(\eta_k)$ bei festem m

(j) $$\left| \frac{m!}{2\pi i} \int\limits_{(-\frac{1}{4})} \frac{e^{x_k \zeta} \{f(\zeta + s) - f(\zeta + s')\}}{\zeta(\zeta+1)(\zeta+2)\ldots(\zeta+m)} d\zeta \right| \leqq \frac{\eta_k}{2}.$$

Ferner kann sodann, wegen

(k) $$\Big| \sum_{\lambda_n \leqq x_k} a_n e^{-\lambda_n \sigma} (e^{-\lambda_n i t} - e^{-\lambda_n i t'}) (1 - e^{\lambda_n - x_k})^m \Big|$$
$$\leqq \sum_{\lambda_n \leqq x_k} |a_n| e^{-\lambda_n (\beta + \frac{1}{2})} |e^{-\lambda_n i (t - t')} - 1|,$$

nach einem klassischen Dirichlet-Kroneckerschen Satze[13]) über diophantische Approximationen die Differenz $d = t - t'$ als ein $d_k = d(\eta_k)$, und zwar

---

[13]) Siehe etwa die einfache Darstellung bei F. Lettenmeyer, Neuer Beweis des allgemeinen Kroneckerschen Approximationssatzes, Proc. of the London Math. Soc. 21 (1922), S. 306–314.

Die hier allein benötigte Formulierung lautet: „*Es seien $y_1, y_2, \ldots, y_r$ beliebige reelle Zahlen. Dann gibt es zu jedem D und jedem $\varepsilon > 0$ ein $d = d(\varepsilon) > D$, so daß für $i = 1, 2, \ldots, r$ zugleich $dy_i - [dy_i] < \varepsilon$ oder $[dy_i] + 1 - dy_i < \varepsilon$ wird.*"

In unserem Falle ist $y_i = \frac{\lambda_i}{2\pi}$ zu setzen.

beliebig groß, so gewählt werden, daß die rechts in (k) auftretenden $|e^{-\lambda_n i(t-t')} - 1|$ sämtlich genügend klein werden, um auch

$$(\mathrm{l}) \qquad \Big|\sum_{\lambda_n \leqq x_k} a_n e^{-\lambda_n \sigma}\left(e^{-\lambda_n i(t'+d_k)} - e^{-\lambda_n i t'}\right)\left(1 - e^{\lambda_n - x_k}\right)^m\Big| \leqq \frac{\eta_k}{2}$$

zu machen, und zwar gleichmäßig für alle unsere s' aus $\mathfrak{K}'$.

Aus (i), (j) und (l) folgt jetzt für alle s' aus $\mathfrak{K}'$

$$(\mathrm{m}) \qquad |f(s' + i d_k) - f(s')| \leqq \eta_k,$$

wenn noch $t' + d_k \geqq 1$ wird, was durch die erfüllbare Auswahl

$$(\mathrm{n}) \qquad d_k \geqq \varrho - t_1 + 1$$

reichlich geschieht.

Setzen wir nun

$$(\mathrm{o}) \qquad f_k(s) = f(s + i d_k),$$

so wird nach (c) und (n) für alle $k = 1, 2, 3, \ldots$ $f_k(s)$ in $\mathfrak{K}$ regulär sein und daselbst

$$(\mathrm{p}) \qquad |f_k(s)| \leqq M$$

gelten, während nach (m) gleichmäßig im konzentrischen $\mathfrak{K}'$

$$(\mathrm{q}) \qquad f_k(s) \to f(s)$$

ist.

Nach einem bekannten Lemma von Stieltjes[14]) folgt nun aus (p) und (q) die Regularität von $f(s)$ im Innern von $\mathfrak{K}$ und also insbesondere im Punkte s_1.

$f(s)$ ist also in der ganzen Halbebene $\sigma > \gamma$ regulär. Wir haben noch zu zeigen, daß $f(s_1) \neq a$ und $\neq b$ wird[15]).

Es bezeichne $\mathfrak{K}_1$ einen zu $\mathfrak{K}$ konzentrischen kleineren Kreis, der aber auch noch s_1 in seinem Innern enthalte. Nach dem Lemma von Stieltjes bleibt dann (q) auch in $\mathfrak{K}_1$ gleichmäßig richtig.

Da $f(s)$ in s_1 regulär ist, kann man offenbar einen gewissen Kreis $\mathfrak{C}$ um s_1, der überdies ganz in $\mathfrak{K}_1$ liegt, angeben, so daß für ein geeignetes $\delta > 0$ auf der Peripherie von $\mathfrak{C}$

$$(\mathrm{r}) \qquad |f(s) - f(s_1)| > 2\delta,$$

---

[14]) Correspondance d'Hermite et de Stieltjes, Paris 1905, **II**, S. 369—370.

Das Lemma lautet: „*Es seien für alle* $k = 1, 2, 3, \ldots$ *die Funktionen* $f_k(z)$ *in* $|z| < 1$ *regulär und daselbst* $|f_k(z)| \leqq M$, *während in* $|z| < \varrho$ *mit* $0 < \varrho < 1$ *gleichmäßig* $f_k(z) \to f(z)$ *gilt. Dann gilt dies letztere sogar gleichmäßig in* $|z| \leqq \vartheta$ *mit irgendeinem* $0 \leqq \vartheta < 1$, *und insbesondere ist also* $f(z)$ *in* $|z| < 1$ *regulär.*"

Über die weitgehenden Verallgemeinerungen dieses Satzes, die wir hier aber nicht benötigen, vgl. die Abhandlg. unter [10]).

[15]) Vgl. die Wennbergsche Dissertation unter [7]), S. 23.

wird. Ebenda wird aber nach (q) für ein genügend großes $k = k(\delta)$

$$(\mathrm{s}) \qquad |f_k(s) - f(s)| < \delta,$$

woraus, immer auf der Peripherie von $\mathfrak{C}$,

$$(\mathrm{t}) \qquad |f_k(s) - f(s_1)| \geqq |f(s) - f(s_1)| - |f_k(s) - f(s)| > \delta > |f_k(s) - f(s)|$$

folgt.

Wäre nun etwa $f(s_1) = a$, also $f_k(s) - f(s_1) = f(s + i d_k) - a$ wegen (7) $\neq 0$ im Innern und auf dem Rande von $\mathfrak{C}$, so folgte ebenda überall aus (t)

$$(\mathrm{u}) \qquad \left|\frac{f_k(s) - f(s)}{f_k(s) - a}\right| < 1.$$

Die in $\mathfrak{C}$ gültige Identität

$$(\mathrm{v}) \qquad f(s) - a = (f_k(s) - a)\left[1 - \frac{f_k(s) - f(s)}{f_k(s) - a}\right]$$

führte nun mit $s = s_1$ zum Widerspruch.

Damit ist der Satz vollständig bewiesen.

Als interessante Folgerung aus Satz 2 wollen wir noch formulieren

Satz 3. *Es sei* (1) *in einer Halbebene* $\sigma > \alpha$ *für irgendein* $\varkappa \geqq 0$ $(\lambda, \varkappa)$-*summierbar. Das hier erklärte* $f(s)$ *sei für* $\sigma \geqq \gamma$, $t \geqq \tau$ *regulär, und es gelte daselbst für ein gewisses k gleichmäßig mit wachsendem t*

$$(9) \qquad f(s) = O(t^k).$$

Ferner sei auf dem Halbstrahle $\sigma = \gamma$, $t \geqq \tau$ *etwa die „einseitige" Beschränkung*

$$(10) \qquad \Re(f(s)) = u(s) \geqq 0$$

erfüllt.

Dann ist $f(s)$ *sogar in der ganzen Halbebene* $\sigma > \gamma$ *regulär und daselbst* (10) *erfüllt.*

Beweis. Ohne Beschränkung der Allgemeinheit sei $\tau = 0$. Ferner setzen wir zum Beweise in (1) $\lambda_0 = 0$ voraus und lassen dafür eventuell $a_0 = 0$ zu.

Wegen

$$(\mathrm{a}) \qquad \lim_{\sigma \to \infty} f(\sigma) = a_0$$

(nach Bemerk. 1) existiert

$$(\mathrm{b}) \qquad m = \operatorname*{Min}_{\sigma \geqq \gamma} u(\sigma).\ ^{16})$$

[16]) Min soll hier als „untere Grenze" verstanden sein.

Aus (10), (b) und (9) folgt nun leicht[17]) durch Betrachtung von $F(s) = e^{-f(s)}$ nach Lindelöf--Phragmén[9]), daß in der Viertelebene $\sigma \geqq \gamma$, $t \geqq 0$

(c) $$u(s) \geqq \operatorname{Min}\{m, 0\} = m_1$$

ist.

Satz 2 lehrt also die Regularität von $f(s)$ in $\sigma > \gamma$ und die Gültigkeit von (c) daselbst. Wenn nun $m_1 = m$ ist, so folgt, da doch $u(s)$ in keinem Punkte $s = \sigma$ mit $\sigma > \gamma$ ein Minimum haben kann (außer wenn $f(s) \equiv a_0$ ist) in jedem Falle

(d) $$m = u(\gamma) \quad \text{oder} \quad = \Re(a_0).$$

Im ersteren Falle muß dann wegen $u(\gamma) \geqq 0$, $m = 0$ also $m_1 = 0$ sein. Aber auch im zweiten Falle folgt, wie ich a. a. O.[18]) gezeigt habe, zunächst, daß $f(s) \equiv a_0$ ist, und also wegen (10), daß $m = 0$ ist.

(c) ist also mit

(e) $$u(s) \geqq 0$$

gleichbedeutend, wie behauptet.

Königsberg i. Pr., den 1. Januar 1924.

---

[17]) Vgl. des Verfassers Zur Theorie der Dirichletschen Reihen, Mathem. Zeitschrift **20** (1924), S. 280—320, Beweis von Satz III.

[18]) Siehe unter [17]), S. 307.

(Eingegangen am 20. 2. 1924.)

Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen.

Von

A. Khintchine in Moskau.

In der vorliegenden Arbeit behandle ich einige Gesetze über die Approximation irrationaler Zahlen mittels rationaler Brüche. Diese Gesetze gelten *fast überall*, d. h. für alle Irrationalzahlen mit Ausnahme einer Menge vom Lebesgueschen Maße Null. Sie gehören also, wenn ich mich so ausdrücken darf, zur metrischen Theorie der Diophantischen Approximationen. Als Stütze gebrauche ich hierbei einige Sätze über Kettenbrüche, welche vielleicht auch ein selbständiges Interesse bieten.

§ 1.

Das geometrische Mittel der Kettenbruchnenner.

Hilfssatz I. *Es ist*

$$\sum_{a_1 a_2 \ldots a_n \geqq g} \frac{1}{(a_1 a_2 \ldots a_n)^2} \leqq 2^n \cdot \frac{\sum_{i=0}^{n-1} \frac{(\lg g)^i}{i!}}{g};$$

dabei ist $g \geqq 1$, *n ganz positiv, und die Summation erstreckt sich über alle ganzen positiven Wertkombinationen der* $a_1, a_2, \ldots, a_n$, *deren Produkt* $\geqq g$ *ist.*

Beweis. Für ein beliebiges Glied der linksstehenden Summe findet man die Abschätzung

$$\prod_{i=1}^{n} \frac{1}{a_i^2} = \prod_{i=1}^{n} \left(1 + \frac{1}{a_i}\right) \frac{1}{a_i(a_i+1)} \leqq 2^n \prod_{i=1}^{n} \frac{1}{a_i(a_i+1)}$$

$$= 2^n \prod_{i=1}^{n} \int_{a_i}^{a_i+1} \frac{dx_i}{x_i^2} = 2^n \int_{a_1}^{a_1+1} \int_{a_2}^{a_2+1} \cdots \int_{a_n}^{a_n+1} \frac{dx_1\, dx_2 \ldots dx_n}{x_1^2 x_2^2 \ldots x_n^2}.$$

Folglich ist

$$\sum_{\Pi a_i \geqq g} \left(\prod_{i=1}^{n} \frac{1}{a_i^2} \right) \leqq 2^n \cdot \int\int \cdots \int \frac{d x_1 \, d x_2 \ldots d x_n}{x_1^2 \, x_2^2 \ldots x_n^2} = 2^n \cdot J_n(g),$$

wo das n-fache Integral über den durch

$$x_1 \, x_2 \ldots x_n \geqq g, \quad x_i \geqq 1 \qquad (i = 1, 2, \ldots, n)$$

bestimmten Bereich zu erstrecken ist. Mittels einer einfachen Substitution findet man leicht

$$J_n(g) = \frac{1}{g} \sum_{i=0}^{n-1} \frac{(\lg g)^i}{i!},$$

womit Hilfssatz I bewiesen ist.

Im folgenden sollen $a_1, a_2, \ldots, a_n, \ldots$ die Kettenbruchnenner der regelmäßigen Kettenbruchentwicklung einer Irrationalzahl bedeuten.

Satz I. *Es ist fast überall*

$$\limsup_{n \to \infty} \sqrt[n]{a_1 \, a_2 \ldots a_n} \leqq e^{e^{\sqrt{2 \lg 2}}}.$$

Beweis. Bekanntlich[1]) bilden die Zahlen, deren n erste Kettenbruchnenner $a_1, a_2, \ldots, a_n$ sind, ein Intervall, dessen Länge

$$\delta < \frac{1}{q^2 \, a_n^2}$$

ist, wo q den Nenner des endlichen Kettenbruchs

$$\frac{p}{q} = [a_1, a_2, \ldots, a_{n-1}]$$

bedeutet. Wegen

$$q \geqq a_1 \, a_2 \ldots a_{n-1}$$

ist also

$$\delta < \prod_{i=1}^{n} \frac{1}{a_i^2}.$$

Folglich ist das Maß der Menge der Zahlen, welche der Bedingung

$$a_1 \, a_2 \ldots a_n \geqq g$$

genügen, nach Hilfssatz I kleiner als

$$2^n \cdot \frac{1}{g} \sum_{i=0}^{n-1} \frac{(\lg g)^i}{i!}.$$

[1]) Siehe z. B. F. Bernstein, Math. Ann. 71 (1912), S. 417.

Ich bezeichne diese Menge mit $E_n(g)$ und setze jetzt $g = e^{An}$, wo $A > e^{\sqrt{2\lg 2}}$ konstant und sonst beliebig sein soll. Ich erhalte somit[2])

$$\mathfrak{M}E_n(g) < 2^n e^{-An} \sum_{i=0}^{n-1} \frac{(An)^i}{i!}.$$

Nun ist, wie leicht einzusehen, jedes Glied der Summe rechts kleiner als $\frac{(An)^n}{n!}$, also

$$\mathfrak{M}E_n(g) < 2^n e^{-An} \cdot n \frac{(An)^n}{n!} < 2^n e^{-An} \cdot n \cdot C_1 \frac{A^n n^n}{n^n e^{-n} \sqrt{2\pi n}}$$
$$< C_2 \sqrt{n}\, e^{-n[A - \lg 2 - \lg A - 1]},$$

wo C_1 und C_2 absolute Konstanten bedeuten.

Nun ist

$$A - 1 - \lg A > \frac{1}{2} \lg^2 A > \lg 2,$$

also

$$\alpha = A - \lg 2 - \lg A - 1 > 0,$$

und folglich ist

$$\mathfrak{M}E_n(g) < C_2 \sqrt{n}\, e^{-\alpha n}$$

in bezug auf n das allgemeine Glied einer konvergenten Reihe.

In bekannter Weise folgert man daraus, daß jede Zahl mit eventueller Ausnahme einer Menge vom Maße Null höchstens einer endlichen Anzahl der Mengen E_n angehören kann und folglich für genügend großes n der Bedingung

$$\prod_{i=1}^{n} a_i < e^{An}$$

Genüge leisten muß. Damit ist aber Satz I bewiesen, denn A darf beliebig nahe an $e^{\sqrt{2\lg 2}}$ genommen werden.

Folgerung. Ich bezeichne mit $\frac{p_n}{q_n}$ den n-ten Näherungsbruch einer irrationalen Zahl. Dann ist $q_1 = a_1$, und für $n > 1$

$$q_n < q_{n-1}(a_n + 1) \leqq 2 a_n q_{n-1},$$

also

$$q_n < 2^n \prod_{i=1}^{n} a_i,$$

also, wegen Satz I, fast überall für genügend großes n und $B > e^{\sqrt{2\lg 2}} + \lg 2$

$$q_n < e^{Bn}.$$

---

[2]) $\mathfrak{M}E$ bedeutet das Maß der Menge E (im Lebesgueschen Sinne).

§ 2.
Ordnung der Annäherung.

Wenn x eine beliebige Irrationalzahl bedeutet, kann man bekanntlich unendlich viele irreduzible Brüche $\frac{p}{q}$ finden von der Beschaffenheit, daß

$$\left|x-\frac{p}{q}\right|<\frac{1}{\sqrt{5}\,q^2}$$

gilt. Größer als $\sqrt{5}$ kann der konstante Faktor im Nenner rechts bekanntlich im allgemeinen nicht gemacht werden[3]). A fortiori kann also auch die Ordnung (in bezug auf q) der Annäherung im allgemeinen nicht vergrößert werden. Nun fragt es sich aber, wie die Sache steht, wenn man auf eine Menge vom Maße Null von Irrationalzahlen verzichten will. Wie weit kann unter dieser Bedingung die Annäherung gefördert werden?

Vollständige Auskunft darüber gibt der folgende

Satz II. *Es sei $f(x)$ eine positive stetige Funktion des positiven Argumentes x, und $x f(x)$ nehme beständig ab. Die Ungleichung*

$$\left|\alpha-\frac{p}{q}\right|<\frac{f(q)}{q}$$

hat fast überall (in bezug auf α) eine unendliche Anzahl von Lösungen in ganzen p, q, wenn $\int^{\infty} f(x)\,dx$ divergiert, und hat fast überall höchstens eine endliche Zahl von Lösungen, wenn $\int^{\infty} f(x)\,dx$ konvergiert.

Beweis. 1. Das Integral $\int^{\infty} f(x)\,dx$, also auch die Reihe $\sum_{n=1}^{\infty} f(n)$, sei konvergent. Ich bezeichne mit M_n die Menge der Zahlen α des Intervalls $(0, 1)$, welche bei geeignet gewähltem ganzen k der Ungleichung

$$\left|\alpha-\frac{k}{n}\right|<\frac{f(n)}{n}$$

genügen. Offenbar ist

$$\mathfrak{M} M_n \leqq 2 f(n),$$

also das allgemeine Glied einer konvergenten Reihe. Bekanntlich gehört daher jede Zahl α der Strecke $(0, 1)$ mit Ausnahme einer Menge vom Maße Null höchstens einer endlichen Anzahl der Mengen M_n an. Damit ist aber die zweite Hälfte von Satz II bewiesen, denn die Beschränkung auf das Intervall $(0, 1)$ ist natürlich unwesentlich.

2. Das Integral $\int^{\infty} f(x)\,dx$, also auch die Reihe $\sum_{n=1}^{\infty} f(n)$ sei divergent. Man setze

$$\varphi(x)=e^{Bx} f(e^{Bx}), \qquad B>e^{\sqrt{2\lg 2}}+\lg 2 \text{ konst.}$$

[3]) Vgl. z. B. Hurwitz, Math. Ann. **39** (1891). S. 279–284.

Die Funktion $\varphi(x)$ ist stetig, positiv und abnehmend, und wir erhalten

$$\lim_{A\to\infty}\int_a^A \varphi(x)\,dx = \lim_{A\to\infty}\frac{1}{B}\int_{e^{Ba}}^{e^{BA}} f(u)\,du = +\infty.$$

Daher folgt nach einem Satze von Herrn F. Bernstein[4]), daß fast überall für unendlich viele i

$$a_{i+1} > \frac{1}{\varphi(i)}$$

ist; daraus folgt aber

$$\left|\alpha - \frac{p_i}{q_i}\right| < \frac{1}{q_i q_{i+1}} < \frac{1}{a_{i+1} q_i^2} < \frac{\varphi(i)}{q_i^2}.$$

Nun ist aber nach Satz I für genügend große i fast überall

$$q_i < e^{Bi}, \quad \text{also} \quad i > \frac{\lg q_i}{B},$$

und folglich

$$\left|\alpha - \frac{p_i}{q_i}\right| < \frac{\varphi\left(\frac{\lg q_i}{B}\right)}{q_i^2} = \frac{f(q_i)}{q_i},$$

womit auch die erste Häfte von Satz II bewiesen ist.

§ 3.

Das arithmetische Mittel der Kettenbruchnenner.

Hilfssatz II. *Es seien $i_1 < i_2 < \ldots < i_r$ ganze positive Zahlen und $g > 0$. $E_r(g)$ sei die Menge der Zahlen*[5]), *welche den Bedingungen $a_{i_k} > g$ $(k = 1, 2, \ldots, r)$ genügen. Dann ist*

$$\mathfrak{M} E_r(g) < \left(\frac{2}{g}\right)^r.$$

Beweis. 1. $r = 1$. Die Zahlen, für welche $a_1, a_2, \ldots, a_{i_1-1}$ bestimmte Werte haben, bilden ein Intervall, dessen Länge ich mit $\delta_{a_1, a_2, \ldots, a_{i_1}-1}$ bezeichne. Die Punkte dieses Intervalls, welche der Bedingung $a_{i_1} > g$ genügen, bilden ein Intervall, dessen Länge nach Herrn F. Bernsteins zitierten Resultaten kleiner als

$$\frac{2}{[g]+1}\,\delta_{a_1, a_2, \ldots, a_{i_1}-1} < \frac{2}{g}\,\delta_{a_1, a_2, \ldots, a_{i_1}-1}.$$

Folglich ist

$$\mathfrak{M} E_1(g) < \frac{2}{g}\sum \delta = \frac{2}{g}.$$

---

[4]) l. c. [1]).

[5]) Wir beschränken uns hier und im folgenden auf das Intervall $(0, 1)$.

2. r sei >1, und für $r-1$ der Satz wahr. Alles bleibt wie für den Fall $r=1$; nur betrachten wir jetzt natürlich ausschließlich solche $\delta_{a_1, a_2, \ldots, a_{i_{r-1}}}$, für welche $a_{i_k} > g$ für $k = 1, 2, \ldots, r-1$ ist. Wir erhalten

$$\mathfrak{M} E_r(g) < \frac{2}{g} \sum \delta < \frac{2}{g} \left(\frac{2}{g}\right)^{r-1} = \left(\frac{2}{g}\right)^r,$$

w. z. b. w.

Hilfssatz III. *$i_1 < i_2 < \ldots < i_n$ seien ganz positiv, $g > 0$, $E_{n,r}(g)$ sei die Menge der Zahlen, für welche mehr als r*[6]*) von den Zahlen $a_{i_1}, a_{i_2}, \ldots, a_{i_n}$ größer als g ausfallen. Dann ist (C absolut konstant)*

$$\mathfrak{M} E_{n,r}(g) < C \left(\frac{2en}{gr}\right)^r.$$

Beweis. G sei eine bestimmte Kombination von $[r]+1$ aus den n Indizes $i_1, i_2, \ldots, i_n$ und E_G bedeute die Menge der Zahlen, für welche alle a_i mit den zur gewählten Kombination gehörigen Indizes größer als g sind. Offenbar gehört jeder Punkt der Menge $E_{n,r}(g)$ wenigstens einer der Mengen E_G; nach Hilfssatz II ist aber

$$\mathfrak{M} E_G < \left(\frac{2}{g}\right)^{[r]+1}.$$

Folglich erhalten wir

$$\mathfrak{M} E_{n,r}(g) \leqq \sum_G \mathfrak{M} E_G < \binom{n}{[r]+1} \cdot \left(\frac{2}{g}\right)^{[r]+1}$$

$$< C' \left(\frac{2en}{g([r]+1)}\right)^{[r]+1} < C' \left(\frac{2en}{rg}\right)^{[r]+1},$$

also auch

$$\mathfrak{M} E_{n,r}(g) < C \left(\frac{2en}{rg}\right)^r$$

(denn im Falle $2en > rg$ können wir ja immer durch $C > 1$ die Behauptung trivial machen), w. z. b w.

Hilfssatz IV. *$\psi(n)$ sei eine positive zunehmende Funktion; dabei sei $\psi(n) < n^2$ für $n \geqq 1$ und die Reihe $\sum_{n=0}^{\infty} \frac{1}{\psi(n)}$ konvergent. Ferner seien m und $i_1 < i_2 < \ldots < i_{2^m}$ ganze positive Zahlen. Endlich sei E_m die Menge der Zahlen, welche folgender Bedingung genügen: Für wenigstens ein ganzes k, $0 < k \leqq m$, sind mehr als*

$$\frac{2e^2 \cdot 2^k}{\psi(k)}$$

---

[6]) $0 \leqq r < n$.

unter den Zahlen $a_{i_1}, a_{i_2}, a_{i_3}, \ldots, a_{i_{s_m}}$ *größer als*

$$2^{m-k}\psi(m).$$

Dann ist, wenn C_1 *eine absolute Konstante bedeutet,*

$$\mathfrak{M}E_m < \frac{C_1}{\psi(m)}.$$

Beweis. Ich bezeichne mit $E_{m,k}$ die Menge der Zahlen, für welche die Bedingung des Satzes für ein bestimmtes k erfüllt ist. Dann folgt aus Hilfssatz III, daß

$$\mathfrak{M}E_{m,k} < C\left(\frac{2e\,2^m\psi(k)}{2e^2 2^k\cdot 2^{m-k}\psi(m)}\right)^{\frac{2e^2\cdot 2^k}{\psi(k)}} = C\left(\frac{\psi(k)}{e\,\psi(m)}\right)^{\frac{2e^2 2^k}{\psi(k)}}$$

ist. Andererseits gehört jeder Punkt der Menge E_m wenigstens einer der Mengen $E_{m,k}$ $(0 < k \leqq m)$ an. Folglich ist

$$\mathfrak{M}E_m < C\sum_{k=1}^{m}\left(\frac{\psi(k)}{e\,\psi(m)}\right)^{\frac{2e^2 2^k}{\psi(k)}}.$$

Die Glieder der rechtsstehenden Summe teilen wir nun in zwei Gruppen, Σ_1 und Σ_2. Dabei nehmen wir in Σ_1 solche Glieder auf, für welche $\sqrt{k}\,\psi(k) \leqq \sqrt{\psi(m)}$ ist. Die Anzahl solcher Glieder sei

$$K-1 \qquad\qquad (k = 1, 2, \ldots, K-1).$$

Es ist leicht einzusehen, daß unter den über $\psi(n)$ gemachten Voraussetzungen immer

$$\frac{2e^2 2^k}{\psi(k)} > 2$$

ist. Folglich ist jedes Glied von Σ_1 kleiner als

$$\left(\frac{\psi(K-1)}{\psi(m)}\right)^2 \leqq \frac{\psi(m)}{(K-1)[\psi(m)]^2} = \frac{1}{(K-1)\,\psi(m)},$$

also

$$(1) \qquad \Sigma_1 < \frac{K-1}{K-1}\cdot\frac{1}{\psi(m)} = \frac{1}{\psi(m)}$$

(K wächst ersichtlich mit m ins Unendliche). Andererseits ist in Σ_2 jedes Glied offenbar nicht größer als

$$e^{-\frac{2e^2 2^K}{\psi(K)}} < e^{-\frac{2e^2\cdot 2^K}{K^2}}.$$

Nun ist

$$\sqrt{K}\cdot K^2 > \sqrt{K}\,\psi(K) > \sqrt{\psi(m)},$$

und folglich

$$K > \sqrt[5]{\psi(m)};$$

ferner ist, wie aus den Voraussetzungen über $\psi(n)$ folgt, für große m-Werte

$$\psi(m) > m,$$

also

$$K > m^{\frac{1}{2}};$$

endlich ist die Anzahl der Glieder von $\sum_2$ nicht größer als m.

Folglich ist für große m-Werte

$$(2) \qquad \sum\nolimits_2 < m e^{-\frac{2e^2 2^K}{K^2}} < m e^{-\frac{2e^2 2^{m^{\frac{1}{2}}}}{m^{\frac{2}{2}}}} < \frac{1}{m^2} < \frac{1}{\psi(m)}.$$

Aus (1) und (2) folgt, daß für große m

$$\mathfrak{M} E_m < C(\sum\nolimits_1 + \sum\nolimits_2) < \frac{2C}{\psi(m)}$$

ist, womit Hilfssatz IV bewiesen ist.

Hilfssatz V. *Unter denselben Voraussetzungen sei F_m die Menge der Zahlen, für welche wenigstens eine der Zahlen $a_{i_1}, a_{i_2}, \ldots, a_{i_{2^m}}$ größer als $2^m \psi(m)$ ist. Dann ist*

$$\mathfrak{M} F_m < \frac{2}{\psi(m)}.$$

Beweis. Folgt sofort aus Hilfssatz II, $r = 1$.

Hilfssatz VI. *Unter denselben Voraussetzungen ist für jeden Punkt, welcher weder der Menge E_m noch der Menge F_m angehört,*

$$s = a_{i_1} + a_{i_2} + \ldots + a_{i_{2^m}} < C_2 \cdot 2^m \psi(m),$$

wo C_2 von m unabhängig ist.

Beweis. Ich bezeichne mit n_k die Anzahl derjenigen unter den Zahlen $a_{i_1}, a_{i_2}, \ldots, a_{i_{2^m}}$, welche zwischen $2^{m-k}\psi(m)$ exkl. und $2^{m-k+1}\psi(m)$ inkl. $(k = 1, 2, \ldots, m)$ enthalten sind. Ferner soll N die Anzahl derer unter diesen Zahlen bezeichnen, welche $\psi(m)$ nicht übersteigen. Da der gewählte Punkt F_m nicht angehört, ist

$$\begin{aligned} s &\leqq N\psi(m) + n_m 2\psi(m) + n_{m-1} 2^2 \psi(m) + \ldots \\ &\quad + n_2 \cdot 2^{m-1} \psi(m) + n_1 \cdot 2^m \psi(m) \\ &= \psi(m)\{N + 2\sum_{k=1}^{m} n_k 2^{m-k}\}. \end{aligned}$$

Nun ist, da der gewählte Punkt nicht der Menge E_m angehört, jedes $n_k \leqq \frac{2e^2 \cdot 2^k}{\psi(k)}$, also

$$s \leqq \psi(m)\left\{N + 4e^2 \cdot 2^m \sum_{k=1}^{m} \frac{1}{\psi(k)}\right\},$$

oder, da $N \leqq 2^m$ ist, wenn noch $D = \sum_{k=1}^{\infty} \frac{1}{\psi(k)}$ gesetzt wird,

$$s < 2^m \psi(m)[1 + 4e^2 D],$$

womit Hilfssatz VI bewiesen ist.

Satz III. *Unter denselben Voraussetzungen über die Funktion $\psi(n)$ ist fast überall*

$$\sigma_n = \frac{1}{n} \sum_{i=1}^{n} a_i = O(\psi(\lg n)).\ ^{1})$$

Beweis. Ich setze $s_n = \sum_{i=1}^{n} a_i$, also $\sigma_n = \frac{s_n}{n}$.

Aus den Hilfssätzen IV und V entnimmt man in bekannter Weise folgendes.

E'_m sei die Menge E_m unter der Voraussetzung $i_k = k$, und E''_m die ähnliche Menge unter der Voraussetzung $i_k = 2^m + k$. Die Mengen F_m des Hilfssatzes V seien für diese beiden Voraussetzungen resp. F'_m und F''_m genannt. Dann ist jeder Punkt, mit Ausnahme einer Menge vom Maße Null, höchstens für eine endliche Anzahl von m-Werten in einer der Mengen E'_m, F'_m, E''_m, F''_m enthalten.

Nun ist wegen Hilfssatz VI für einen solchen Punkt, wenn L eine nur von der Natur der Funktion $\psi(n)$ abhängende Konstante bedeutet, für große m-Werte erstens

$$s_{2^m} < L\, 2^m \psi(m)$$

und zweitens

$$s_{2^{m+1}} - s_{2^m} < L\, 2^m \psi(m),$$

also, wenn

$$2^m \leqq n < 2^{m+1}$$

ist,

$$s_n < s_{2^{m+1}} = s_{2^m} + (s_{2^{m+1}} - s_{2^m}) < 2L\, 2^m \psi(m) \leqq 2Ln\psi(\lg n),$$

womit Satz III bewiesen ist.

Korollar. *Es sei $\varphi(n)$ eine positive Funktion, so daß $\frac{\varphi(n)}{n}$ beständig zunimmt und $\sum_{n=1}^{\infty} \frac{1}{\varphi(n)}$ konvergiert. Dann ist fast überall*

$$s_n = O(\varphi(n)).$$

---

[1]) Der Bequemlichkeit halber nehme ich die Logarithmen zur Basis 2.

Beweis. Ich setze $\psi(n) = \frac{1}{2^n}\varphi(2^n)$. Dann ist $\psi(n)$ eine zunehmende positive Funktion, und bekanntlich ist

$$\sum \frac{1}{\psi(n)} = \sum \frac{2^n}{\varphi(2^n)}$$

konvergent. Die Annahme $\psi(n) < n^2$, die wir in Hilfssatz IV gemacht haben, kann natürlich bei Satz III ohne weiteres weggelassen werden. Wir können somit Satz III anwenden, und erhalten fast überall

$$s_n = \sum_{i=1}^{n} a_i = O(n\,\psi(\lg n)) = O(\varphi(n)),$$

w. z. b. w.

Das bewiesene Korollar ist eine wesentliche Verschärfung des Hauptsatzes der zitierten Betrachtungen Herrn F. Bernsteins. Dieser Hauptsatz behauptet nämlich, daß unter der Voraussetzung der Konvergenz von $\sum \frac{1}{\varphi(n)}$ fast überall

$$a_n = O(\varphi(n))$$

ist. Unser Korollar zeigt, daß auch

$$s_n = O(\varphi(n))$$

fast überall erfüllt ist; und zwar braucht dazu $\varphi(n)$ nicht schneller zu wachsen, als es der Bernsteinsche Satz verlangt. Nur ist die Forderung an die Regelmäßigkeit des Wachstums hier eine größere, denn $\frac{\varphi(n)}{n}$ muß immer zunehmen.

§ 4.

Die Hardy-Littlewoodschen Summen.

Die grundlegenden Arbeiten der Herren Hardy und Littlewood über die Summen der Gestalt

$$S_n(x) = \sum_{k=1}^{n} \left(kx - [kx] - \tfrac{1}{2}\right),$$

die mit dem Problem der Anzahl der Gitterpunkte in Polygonen eng zusammenhängen, sind mir leider nicht zugänglich geworden. Einiges habe ich darüber veröffentlicht*). Dank einer liebenswürdigen Sendung von

*) Bulletin de l'Institut Polytechnique à Jvanowo-Wosniessensk Nr. 5, Janvier 1922, p. 27—41. Vgl. auch meine während des Druckes erschienene Abhandlung „Ein Satz über Kettenbrüche usw.", Mathematische Zeitschrift 18 (1923), S. 289—306.

Herrn Ostrowski habe ich erfahren, daß ihm einige von meinen Resultaten schon 1921 bekannt und von ihm publiziert waren[9]).

Insbesondere reichen die Formeln der Seite 80 der zitierten Arbeit von Herrn Ostrowski (sowie auch die Formeln der Seiten 33—34 meiner Abhandlung) vollständig aus, um folgenden Satz sehr leicht zu beweisen:

Satz IV. *Es sei $\varphi(n)$ eine positive Funktion und $\frac{\varphi(n)}{n}$ beständig zunehmend. Die Abschätzung*

$$S_n(x) = O(\varphi(\lg n))$$

ist fast überall richtig oder fast überall falsch, je nachdem die Reihe $\sum \frac{1}{\varphi(n)}$ konvergent oder divergent ist.

Die Stütze beim Beweise bilden Satz I und Satz III mit ihren Folgerungen.

Endlich sei noch bemerkt, daß mit der von Herrn Ostrowski und mir entwickelten Methode die bewiesenen kettenbruchtheoretischen Abschätzungen noch auf viele andere metrischen Probleme der Theorie der Diophantischen Approximationen Anwendung finden können. So erhält man zum Beispiel ein Analogon zu Satz IV über die Häufigkeit der Zahlen $kx - [kx]$ in bestimmten Intervallen u. dgl. m.

Moskau, den 7. IX. 1923.

---

[9]) Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität **1** (1921), Heft 1, S. 77—98.

(Eingegangen am 15. 9. 1923.)

Zur Bestimmung der metazyklischen Minimalbasis von Primzahlgrad.

Von

S. Breuer in Karlsruhe i. B.

Die vorliegende Arbeit bietet einen Beitrag zur Bestimmung einer *Minimalbasis für den Invariantenkörper der metazyklischen Permutationsgruppen von Primzahlgrad* p, d. h. einer aus p algebraisch unabhängigen Funktionen bestehenden Basis, durch die alle Funktionen des Körpers sich rational darstellen lassen[1]). Eine solche Basis wurde dem Verfasser für die Grade $p = 5$ und $p = 7$ von Fräulein Noether mitgeteilt; die Noetherschen Formeln werden im § 2 in sachgemäßer Bezeichnung wiedergegeben, mit unwesentlichen Änderungen, die den allmählichen Aufstieg vom metazyklischen Invariantenkörper zu den darüber liegenden deutlicher hervortreten lassen sollen. Zur Basisbestimmung wird nämlich die Tatsache benutzt, daß der Körper aller rationalen Funktionen der Unbestimmten ein Galoisscher Körper über dem Invariantenkörper der metazyklischen Gruppe ist, und daß diese letztere seine Galoissche Gruppe wird. Das Verfahren zerfällt nun in zwei Schritte: Zuerst werden, in den genannten speziellen Fällen, drei bzw. vier Basisfunktionen angegeben — von denen übrigens jeweils die ersten beiden schon länger bekannt sind —, welche der Minimalbasis des Körpers aller rationalen Funktionen der Unbestimmten ebenso angehören können wie der gesuchten Minimalbasis des vollmetazyklischen Körpers, wie auch der jedes Galoisschen Zwischenkörpers. Durch ihre Einführung wird die Aufgabe darauf reduziert, die Minimalbasis für den *zyklischen* Körper von nur zwei bzw. drei Unbestimmten zu finden. Dieser erste Schritt wird im folgenden für den Primzahlgrad p allgemein durchgeführt: Im § 1 werden nämlich $\frac{p+1}{2}$ solche Basisfunktionen an-

[1]) Vgl. E. Noether, Gleichungen mit vorgeschriebener Gruppe, Math. Annalen 78.

gegeben, welche ebenfalls allen in Betracht kommenden Minimalbasen angehören können und die Aufgabe der Aufstellung der vollmetazyklischen Basis auf das freilich noch ungelöste Problem zurückführen, die Minimalbasis für einen zyklischen Körper $\frac{p-1}{2}$-ten Grades zu finden; und zwar handelt es sich dabei um eine zyklische Gruppe von Cremona-Transformationen — birationalen Transformationen unabhängiger Variablen —, die nicht notwendig linear zu sein brauchen. Im speziellen Falle $p=5$ und $p=7$ handelt es sich jedoch um die zyklische *Permutationsgruppe* von $\frac{p-1}{2}$, d. h. zwei bzw. drei Unbestimmten selbst. Es kann daher in diesen beiden Fällen, die weiterhin (§ 2) ausschließlich behandelt werden, die Aufgabe restlos gelöst werden. Es werden nämlich, ausgehend von dem Körper der Unbestimmten selbst — der also wie alle jetzt noch zu behandelnden Körper nur noch zwei bzw. drei Unbestimmte enthält — nacheinander für alle Galoisschen Zwischenkörper Minimalbasen eingeführt, deren Elemente dann unter der metazyklischen Gruppe Substitutionen einer Gruppe erfahren, die der entsprechenden Faktorgruppe isomorph ist. Es brauchen also nur noch für die zu den betreffenden zyklischen Gruppen gehörigen Invariantenkörper die Minimalbasen gebildet zu werden, was leicht auszuführen ist. Diese Minimalbasen nun hatten in der ursprünglichen, eingangs erwähnten Mitteilung von Fräulein Noether — mit Ausnahme der vollmetazyklischen Basen selbst — noch Koeffizienten aus dem Körper der fünften bzw. siebenten Einheitswurzel. Der Verfasser konnte durch „Adjunktion natürlicher den Zwischenkörpern entnommener Funktionen" zu den vollmetazyklischen Basen, und anschließende Rechnung auch zu jedem dieser Zwischenkörper eine Minimalbasis mit *rationalen* Koeffizienten bestimmen. Auf Grund seiner Mitteilung bemerkte Fräulein Noether, daß ihre ursprüngliche Darstellung sich so umgestalten läßt, daß die erwähnte Rechnung überflüssig wird und unmittelbar für alle Zwischenkörper rationale Minimalbasen gewonnen werden (vgl. auch Fußnote 11). Die Kenntnis *rationaler* Minimalbasen auch für die Zwischenkörper ermöglicht es, über beliebigem (endlichen) algebraischen Zahlkörper als Grundkörper alle Galoisschen Körper der entsprechenden Gruppe direkt zu bilden und ist auch deshalb von Bedeutung, weil mit ihr zugleich die Möglichkeit erwiesen ist, die sämtlichen auflösbaren Gleichungen 5. und 7. Grades, *nach ihren Galoisschen Gruppen getrennt*, durch unabhängige Parameter darzustellen. Es ergibt sich eine sehr einfache „Wurzeldarstellung" für die gewählte, von $x_0, x_1, \ldots, x_4$ bzw. $x_0, x_1, \ldots, x_6$ freilich verschiedene, Basis des Körpers aller rationalen Funktionen der Unbestimmten x_ν, woraus unmittelbar die oben erwähnte Bildung der Galoisschen Körper folgt. Dagegen spielen die symmetrischen Funktionen

der x_ν gar keine ausgezeichnete Rolle unter den metazyklischen Funktionen überhaupt; ihre Darstellung durch die Funktionen der Minimalbasis, die zu der vorher erwähnten Parameterdarstellung der metazyklischen Gleichungen führt[1]), wird wenig übersichtlich, es ist daher von dieser abgesehen worden. Es werden jedoch im § 3 die notwendigsten Ergänzungsformeln zu der Noetherschen Minimalbasis für $p = 5$ gegeben, die zur Parameterdarstellung gewisser *spezieller* Gleichungen dieses Grades noch erforderlich sind. Die Darstellung der Gleichungskoeffizienten selbst konnte um so eher unterlassen werden, als Herr Wäisälä[2]) für die metazyklischen Gleichungen 5. Grades eine solche Parameterdarstellung bereits gegeben hat. Die dort auftretenden Parameter werden ebenfalls als Funktionen einer Minimalbasis erkannt. Sie sind speziell dem *Gleichungs*-Problem angepaßt und gestatten daher eine sehr einfache Darstellung der symmetrischen Funktionen und damit der Gleichungskoeffizienten. An sie anknüpfend werden für die über dem metazyklischen Körper liegenden Körper ebenfalls Minimalbasen bestimmt und damit auch für diese Darstellung der Einblick in die Struktur des Körpers eröffnet. Die Minimalbasen für den halbmetazyklischen und den zyklischen Körper sind, ebenso wie die ursprünglichen entsprechenden Noetherschen Basen, nicht frei von fünften Einheitswurzeln. Während man jedoch dort, wie erwähnt, infolge des gruppentheoretischen Standpunktes unmittelbar auch zu Minimalbasen mit *rationalen* Koeffizienten gelangen konnte, ist dies hier nicht der Fall und man müßte sich der vorerwähnten Rechnung des Verfassers bedienen, um von der vollmetazyklischen Basis des Herrn Wäisälä ausgehend auch für die Oberkörper Minimalbasen mit *rationalen* Koeffizienten zu finden. Während wir ferner im § 1 einen wesentlichen Teil des Noetherschen Verfahrens für allgemeinen Primzahlgrad durchführen konnten, zeigt es sich, daß die Anwendung des Verfahrens des Herrn Wäisälä auch nur auf Gleichungen 7. Grades schon rein rechnerisch auf fast unüberwindliche Schwierigkeiten stößt.

§ 1.

Im folgenden bezeichne $p = 2n + 1$ eine Primzahl, ε eine beliebig, aber fest gewählte primitive p-te Einheitswurzel, g eine feste primitive Kongruenzwurzel von p. Die auftretenden Indizes sollen folgende Werte durchlaufen: $j = 0, 1, \ldots, n-1$; $t = 1, 2, \ldots, n-1$; $\mu = 0, 1, \ldots, p-2$; $\nu = 0, 1, \ldots, p-1$. Mit x_ν bezeichnen wir p Unbestimmte, und es bedeute:

$$(1) \qquad (\varepsilon, x) = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \ldots + \varepsilon^{p-1} x_{p-1}.$$

---

[2]) Über die algebraisch auflösbaren Gleichungen 5. Grades. Diss. Helsingfors 1916.

Ist weiter[3])

$$(2)\qquad \frac{1}{p}(1,x) = \frac{1}{p}\sum x_\nu = \varphi_0(x) \quad \text{und} \quad \frac{1}{p}(\varepsilon^{g^\mu}, x) = k_\mu,$$

so wird

$$(3)\qquad x_\nu = \varepsilon^{-\nu}k_0 + \varepsilon^{-g\cdot\nu}k_1 + \varepsilon^{-g^2\cdot\nu}k_2 + \dots + \varepsilon^{-g^{p-2}\cdot\nu}k_{p-2} + \varphi_0(x).$$

Es sei ferner

$$(4)\qquad \frac{k_{n+j}}{k_j} = q_j, \quad \frac{k_j}{k_0} = r_j; \quad \text{entsprechend} \quad q_{n+j} = \frac{1}{q_j},\text{[4]}) \quad r_{n+j} = \frac{k_{n+j}}{k_0} = q_j r_j.$$

Dabei ist also $r_0 = r_{2n} = 1$, $r_n = q_0$, und nur die r_t sind von den q_j verschiedene Unbestimmte, d. h. algebraisch unabhängige Funktionen; es ist, wegen $r_{2n+j} = r_j = \frac{r_{n+j}}{q_j}$, auch $r_{2n+j} = q_{n+j} r_{n+j}$, was wir für späteren Gebrauch anmerken. Wir bezeichnen ferner die erzeugenden Elemente der vollen linearen Permutationsgruppe der x_ν mit

$$(5)\qquad S = \binom{\nu}{\nu+1} = (x_0, x_1, \dots, x_{p-1}), \quad T = \binom{\nu}{g\nu} = (x_1, x_g, x_{g^2}, \dots, x_{g^{p-2}}).$$

Unter „zyklischen oder metazyklischen Funktionen der Größen k_μ, q_j, r_t" sind im folgenden stets Ausdrücke verstanden, welche die Vertauschungen S bzw. S und T (5) dulden. Die „einförmige" ganze rationale Funktion der k_μ

$$(6)\qquad v_0 = k_0^{\lambda_0} k_1^{\lambda_1} \dots k_{p-2}^{\lambda_{p-2}}$$

ist dann und nur dann zyklisch, wenn

$$(7)\qquad \lambda_0 + g^1\lambda_1 + g^2\lambda_2 + \dots + g^{p-2}\lambda_{p-2} \equiv 0 \pmod{p},$$

da k_μ (2) bei Anwendung von S (5) mit ε^{-g^μ} multipliziert wird. Bei Anwendung von T (5) geht dagegen k_μ in $k_{\mu-1}$ über, d. h. die k_μ werden in der Reihenfolge $(k_{p-2}, k_{p-3}, \dots, k_2, k_1, k_0)$ vertauscht. *Ersetzt* man andererseits die Einheitswurzel ε durch ε^g — es handelt sich *nicht* um ein *gegenseitiges* Vertauschen dieser beiden Größen —, so werden die k_μ ersichtlich in der umgekehrten Reihenfolge vertauscht wie vorher durch T. Die Summe — ebenso natürlich *jede* zyklische Funktion — der aus v_0 (6) durch zyklische Verschiebung der k_μ oder der λ_μ hervorgehenden Glieder v_μ

$$(8)\qquad \sum v_\mu = [v_0] = [k_0^{\lambda_0} k_1^{\lambda_1} \dots k_{p-2}^{\lambda_{p-2}}]$$

---

[3]) Vgl. Wäisälä a. a. O. sowie Breuer, Das Abelsche Gleichungsproblem bei Euler, Jahresbericht der Deutschen Mathematiker-Vereinigung 1922, S. 165ff.

[4]) Zunächst ist $q_{n+j} = \frac{k_{2n+j}}{k_{n+j}}$; die Indizes von k, q, r sind aber überall auf den kleinsten positiven Rest (mod $2n$) zu reduzieren.

ist daher eine vollmetazyklische Funktion der k_μ bzw. x_ν, sie ist überdies, im Gegensatz zu v_0 selbst, auch als Funktion der x_ν rational, d. h. frei von ε. Denn da sie sich nicht ändert, wenn ε durch ε^g ersetzt wird, gestatten ihre Koeffizienten die Gruppe der Kreisteilungsgleichung $\frac{x^p-1}{x-1}=0$, sind also rational. Endlich kann jede rationale metazyklische Funktion der x_ν aus φ_0 (2) und Ausdrücken $[v_\mu]$ (8) zusammengesetzt werden[5]). Unter den v_μ kann als „Leitglied“ v_0 irgendein beliebiges herausgegriffen werden. Zwei Leitglieder v_0, v_0' bezeichnen wir als unabhängig, wenn sie nicht durch zyklische Verschiebung der k_μ oder der λ_μ auseinander hervorgehen, d. h. wenn $[v_0] \neq [v_0']$.

Nun seien unter $f_m([v_0])$, $g_m([v_0])$ ganze rationale, aus einem oder mehreren Ausdrücken (8) zusammengesetzte, homogene Funktionen m-ten Grades der k_μ und damit der x_ν verstanden, wobei also $\lambda_0+\lambda_1+\ldots+\lambda_{p-2}=m$. Dann ist

$$\varphi_1(x)=\frac{f_{m_1+1}([v_0])}{g_{m_1}([v_0])} \tag{9}$$

eine rationale metazyklische homogene Funktion 1. Grades der x_ν, und

$$\varphi_{t+1}(x)=\frac{f_{m_{t+1}}([v_0])}{g_{m_{t+1}}([v_0])} \tag{10}$$

sind ebensolche Funktionen 0-ten Grades, die auch als Funktionen der Größen q_j und r_t (4) statt der k_μ geschrieben werden können. Da also jede rationale metazyklische Funktion der x_ν mit Hilfe von (3) durch $\varphi_0(x)$ (2), die beliebig zu wählende Funktion $\varphi_1(x)$ (9) und metazyklische Funktionen der Größen (4) ausgedrückt werden kann, so kann eine metazyklische Minimalbasis der x_ν aus φ_0, φ_1 und einer nur noch $p-2$ Funktionen enthaltenden Basis der metazyklischen Funktionen der q_j, r_t (4) zusammengesetzt werden. Diese Reduktion ist schon lange bekannt[6]). Wiederholt sei bemerkt, daß dabei ε nicht als dem zugrunde gelegten Rationalitätsbereich adjungiert gedacht ist und auch nur scheinbar in die Darstellung eingeht.

Eine wesentliche Reduktion erfährt das Basisproblem nun durch den folgenden Umstand:

Es ist stets möglich, $n-1$ *metazyklische Funktionen* (10) *anzugeben, mit deren Hilfe die* r_t *rational durch die* q_j *ausgedrückt werden können.*

Dazu genügt es, $n-1$ Funktionen (10) anzugeben, die gebrochene *lineare* Funktionen mit nichtverschwindender Determinante der „Unbe-

[5]) Vgl. Weber, Lehrbuch der Algebra, 2. Aufl. (1898), I, §§ 191, 192.

[6]) Vgl. E. Noether a. a. O. S. 227 sowie Breuer, Über die irreduktiblen auflösbaren trinomischen Gleichungen 5. Grades, Diss. Frankfurt 1918, S. 17 ff.

kannten“ r_t sind, da diese alsdann aus dem „linearen Gleichungssystem“ (10) berechnet werden können. [Die „Unbestimmten“ x_ν sollen nur so spezialisiert werden, daß die Determinante dieses Systems und alle Nenner $g_{m_{t+1}}([v_0])$ von Null verschieden bleiben.] Im *speziellen* Fall wird man die $\varphi_{t+1}(x)$ durch Tabulierung der überhaupt möglichen Funktionen $[v_0]$ niederer Grade je nach dem beabsichtigten Zweck möglichst einfach bilden. Zum *allgemeinen* Beweis unserer Behauptung genügt die Angabe irgendeines Systems (10) (von dem die im § 2 anzugebenden Systeme dann freilich wesentlich unterschieden sind). Sei eine Funktion (10) etwa

$$\varphi_2(x) = \frac{[k_0^{\lambda_0} k_1^{\lambda_1} \dots k_{p-2}^{\lambda_{p-2}}]}{[k_0^{l_0} k_1^{l_1} \dots k_{p-2}^{l_{p-2}}]},$$

wo $\sum \lambda_\mu = \sum l_\mu = m_2$, so wird bei Kürzung durch $k_0^{m_2}$ nach (4)

$$\varphi_2(x) = \frac{[r_1^{\lambda_1} r_2^{\lambda_2} \dots r_{n-1}^{\lambda_{n-1}} q_0^{\lambda_n} (q_1 r_1)^{\lambda_{n+1}} (q_2 r_2)^{\lambda_{n+2}} \dots (q_{n-1} r_{n-1})^{\lambda_{2n-1}}]}{[r_1^{l_1} r_2^{l_2} \dots r_{n-1}^{l_{n-1}} q_0^{l_n} (q_1 r_1)^{l_{n+1}} (q_2 r_2)^{l_{n+2}} \dots (q_{n-1} r_{n-1})^{l_{2n-1}}]}$$

$$= \frac{[q_0^{\lambda_n} q_1^{\lambda_{n+1}} q_2^{\lambda_{n+2}} \dots q_{n-1}^{\lambda_{2n-1}} \cdot r_1^{\lambda_1 + \lambda_{n+1}} r_2^{\lambda_2 + \lambda_{n+2}} \dots r_{n-1}^{\lambda_{n-1} + \lambda_{2n-1}}]}{[q_0^{l_n} q_1^{l_{n+1}} q_2^{l_{n+2}} \dots q_{n-1}^{l_{2n-1}} \cdot r_1^{l_1 + l_{n+1}} r_2^{l_2 + l_{n+2}} \dots r_{n-1}^{l_{n-1} + l_{2n-1}}]}.$$

Die eckigen Klammern in Zähler und Nenner bedeuten die Summation über alle Glieder, die aus dem „Leitglied“ durch zyklische Vertauschung der λ_μ bzw. l_μ entstehen. Es ist also $\varphi_2(x)$ *linear* in sämtlichen r_t, wenn die Zahlen: $\lambda_0 + \lambda_n$, $\lambda_t + \lambda_{n+t}$, $l_0 + l_n$, $l_t + l_{n+t}$ nicht *sämtlich* einander gleich sind, und nur die beiden Werte $2u$ oder $2u+1$, letzteren nur einmal, bzw. $2u$, $2u-1$, für irgendein $u \geqq 0$ annehmen, da man alsdann durch $(r_1 r_2 \dots r_{n-1})^{2u}$ kürzen kann. Das gleiche gilt für alle $\varphi_{t+1}(x)$. Ein einfaches System (10) gewinnen wir folgendermaßen. Das Leitglied

(11) $\qquad v_{0,1} = (k_0 k_1 \dots k_{n-1})^\pi k_0, \quad$ wo $\quad p^2 + n(1-g) = \pi$

gesetzt ist [7]), genügt sowohl der Bedingung (7) wie der soeben für die λ_μ aufgestellten Bedingung. Denn

$$\pi(g^0 + g^1 + \dots + g^{n-1}) + g^0 \equiv n(1-g)\frac{g^n - 1}{g-1} + 1 \equiv 2n + 1 \equiv 0 \pmod p,$$

da $g^n \equiv -1 \pmod p$; ferner ist $\lambda_t + \lambda_{n+t} = \pi$, $\lambda_0 + \lambda_n = \pi + 1$. Die beiden erwähnten Bedingungen werden nun auch von den Leitgliedern

(12) $\qquad v_{0,t+1} = (k_0 k_1 \dots k_{n-1})^\pi k_0 \left(\frac{k_n}{k_0}\right)^{tp},$

---

[7]) Ebenso könnte man auch $\pi = fp + n(1-g)$ setzen, wo f eine beliebige ganze Zahl ist, wenn nur $\pi + 1 \geqq (n-1)p$ ist, damit die Größen $v_{0,t+1}$ (12) ganze Funktionen bleiben.

die ganze Funktionen auch von k_0 sind, erfüllt. Denn der Faktor $\left(\frac{k_n}{k_0}\right)^{t\nu}$ genügt der Bedingung (7) und ändert nicht die Zahlen $\lambda_j + \lambda_{n+j}$. Die Leitglieder (11) und (12) sind ferner „unabhängig“ in dem oben definierten Sinne, wie eine einfache Betrachtung lehrt. Daher bilden die Funktionen

$$(13) \qquad \varphi_{t+1}(x) = \frac{[v_{0,t+1}]}{[v_{0,1}]}$$

ein System der verlangten Art (10), sobald das Nichtverschwinden der Determinante gezeigt ist. Führt man in der ausführlich geschriebenen Formel (13) mit Kürzung durch $k_0^{n\pi+1}$ die Größen (4) ein und kürzt erneut durch $(r_1 r_2 \ldots r_{n-1})^\pi$, so erhält man

$$(14) \qquad \varphi_{t+1}(x) = \frac{Z_{t+1}}{N},$$

wobei

$$(14a) \qquad \begin{aligned} N = r_0 + q_0^\pi r_1 + (q_0 q_1)^\pi r_2 + \cdots + (q_0 q_1 \cdots q_{n-1})^\pi r_n \\ + (q_1 q_2 \cdots q_{n-1})^\pi r_{n+1} + (q_2 q_3 \cdots q_{n-1})^\pi r_{n+2} + \cdots \\ + (q_{n-2} q_{n-1})^\pi r_{2n-2} + q_{n-1}^\pi r_{2n-1}, \end{aligned}$$

und wobei Z_{t+1} aus N entsteht, indem der Koeffizient von r_μ mit $q_\mu^{t\nu}$ multipliziert wird. Da nun nach (4) $r_{n+j} = q_j r_j$, $r_0 = 1$, so wird

$$(14b) \qquad \begin{aligned} N = [1 + q_0(q_0 q_1 \cdots q_{n-1})^\pi] + [q_0^\pi + q_1(q_1 q_2 \cdots q_{n-1})^\pi] r_1 \\ + [(q_0 q_1)^\pi + q_2(q_2 q_3 \cdots q_{n-1})^\pi] r_2 + \cdots \\ + [(q_0 q_1 \cdots q_{j-1})^\pi + q_j(q_j q_{j+1} \cdots q_{n-1})^\pi] r_j + \cdots \\ + [(q_0 q_1 \cdots q_{n-3})^\pi + q_{n-2}(q_{n-2} q_{n-1})^\pi] r_{n-2} \\ + [(q_0 q_1 \cdots q_{n-2})^\pi + q_{n-1} \cdot q_{n-1}^\pi] r_{n-1}, \end{aligned}$$

und Z_{t+1} entsteht aus N, indem in der j-ten eckigen Klammer das erste Glied mit $q_j^{t\nu}$, das zweite mit $q_{n+j}^{t\nu}$, d. h. mit $q_j^{-t\nu}$ multipliziert wird. Bezeichnen wir nun die Koeffizienten der r_j im Polynom N mit a_{0j}, in den Polynomen Z_{t+1} mit a_{tj}, so können wir allgemein schreiben

$$(15) \qquad a_{hj} = (q_0 q_1 \cdots q_{j-1})^\pi q_j^{h\nu} + q_j (q_j q_{j+1} \cdots q_{n-1})^\pi q_j^{-h\nu}$$
$$(h, j = 0, 1, \ldots, n-1).$$

Zum Beweise unserer Behauptung, daß *die r_t mit Hilfe der $n-1$ metazyklischen Funktionen φ_{t+1} (13) rational durch die q_j ausgedrückt werden können*, haben wir nunmehr noch zu zeigen, daß die Determinante D des Gleichungssystems (14), d. h. die Determinante $\{a_{h,j}\}$ nicht *identisch* verschwindet, da die *speziellen* Werte (der x_ν), für die sie verschwindet, schon früher ausgeschlossen worden sind. Hierzu genügt es, da die q_j

algebraisch unabhängige Funktionen, also Unbestimmte sind, *ein* Wertesystem der q_j anzugeben, für welches $D = \{a_{h,j}\} \neq 0$ ist. Wir wählen für die q_j nun n verschiedene, paarweise zueinander reziproke Zahlen — es sei stets $q_{2i} q_{2i+1} = 1$ —, die sonst nur der Bedingung unterliegen, daß $q_j^{2\pi} \neq 1$ sei; nur wenn n ungerade ist, sei $q_{n-1} = 1$ gesetzt. Dann wird in leichtverständlicher Bezeichnung

$$(15\,\mathrm{a}) \qquad a_{h,2i} = q_{2i}^{hp} + q_{2i}^{1-hp}; \qquad a_{h,2i+1} = q_{2i}^{\pi-hp} + q_{2i}^{-\pi-1+hp}.$$

Für ungerades n wird weiter $a_{h,n-1} = 2$. Zieht man nun das $q_{2i}^{-\pi-1}$-fache der $2i$-ten Spalte von der $2i+1$-ten ab, so wird

$$\bar{a}_{h,2i+1} = q_{2i}^{-hp}\,(q_{2i}^{\pi} - q_{2i}^{-\pi}) \neq 0$$

nach der Voraussetzung, daß $q_j^{2\pi} \neq 1$. Zieht man weiter das $\dfrac{q_{2i}}{q_{2i}^{\pi} - q_{2i}^{-\pi}}$-fache der so transformierten $2i+1$-ten Spalte von der $2i$-ten ab, so wird $\bar{a}_{h,2i} = q_{2i}^{hp}$. Dividiert man also D durch die sämtlichen Faktoren $q_{2i}^{\pi} - q_{2i}^{-\pi}$ und allenfalls, für ungerades n, noch durch 2, so erhält man die bekannte Darstellung des Differenzenproduktes der q_j, das nach unserer Wahl der q_j von Null verschieden ist. Für diese q_j ist daher $D \neq 0$, w. z. b. w.

Die gesuchte Minimalbasis kann also jetzt zusammengesetzt werden aus den $n+1$ Funktionen φ_0 (2), φ_1 (9), φ_{t+1} (13) — die rationale Funktionen der x_ν sind mit Koeffizienten aus dem absoluten Rationalitätsbereich — und einer aus nur noch n Funktionen bestehenden Basis der metazyklischen Funktionen der q_j. Denn aus $\varphi_0, \varphi_1, \varphi_{t+1}, q_j$ geht derselbe Körper hervor wie aus $\varphi_0, \varphi_1, q_j, r_t, r_{n+t}$; er ist vom Grade $p(p-1)$ über der gesuchten Minimalbasis.

Für die *zyklischen* Funktionen der q_j bilden nun die Größen

$$(16) \qquad q_0^p, \quad q_t q_0^{-g^t}$$

eine freilich von ε nicht freie Basis, denn nach Adjunktion der einen p-ten Wurzel q_0 sind alle q_j bekannt. Durch T^n (5) gehen sie in ihre reziproken Werte über. Die mit ihnen birational zusammenhängenden und also ebenfalls eine zyklische Basis bildenden Größen

$$(16\,\mathrm{a}) \qquad \frac{q_0^p - 1}{q_0^p + 1}, \quad \frac{q_t q_0^{-g^t} - 1}{q_t q_0^{-g^t} + 1}$$

ändern also bei T^n nur ihr Vorzeichen. Daher bilden die n Größen

$$(17) \qquad \left(\frac{q_0^p - 1}{q_0^p + 1}\right)^2, \quad \frac{q_0^p - 1}{q_0^p + 1} \cdot \frac{q_t q_0^{-g^t} - 1}{q_t q_0^{-g^t} + 1}$$

eine Basis der die Vertauschungen S und T^n duldenden Funktionen[8]) der q_j, bzw. zusammen mit den $\varphi_0, \varphi_1, \varphi_{t+1}$ eine solche in bezug auf die x_ν selbst. In gewissen Fällen wird es möglich sein, die n Größen (17) durch solche zu ersetzen, die durch T in einfacherer Weise geändert, möglicherweise nur zyklisch vertauscht werden. Zur Vervollständigung der gesuchten Minimalbesis bedarf man alsdann nur noch einer solchen für den Körper der zyklischen Funktionen dieser n Unbekannten. Im *allgemeinen* Fall, der hier nicht weiter behandelt werden soll, erfahren die Größen (17) durch T eine Gruppe von Cremona-Transformationen, die der im speziellen Fall auftretenden zyklischen Permutationsgruppe von n Unbestimmten isomorph ist[9]). Diese selbst ist wieder nichts anderes als die letzte Faktorgruppe in unserer Kompositionsreihe der vollmetazyklischen Gruppe.

§ 2.

Im Falle $p = 5$ wählen wir $g = 2$. Nach Tabulierung der die Bedingung (7) erfüllenden Funktionen v_0 (6) der ersten Grade[10]) findet man als einfachste Ausdrücke:

$$(9\text{a}) \qquad \varphi_1 = \frac{k_0^2 k_3 + k_1^2 k_0 + k_2^2 k_1 + k_3^2 k_2}{k_0 k_2 + k_1 k_3} = k_0 \frac{q_1 r_1 + r_1^2 + q_0^2 r_1 + q_0 q_1^2 r_1^2}{q_0 + q_1 r_1^2},$$

$$(13\text{a}) \qquad \varphi_2 = \frac{k_0 k_1 k_2 k_3 (k_0^2 k_3 + k_1^2 k_0 + k_2^2 k_1 + k_3^2 k_2)}{k_0^4 k_1^3 + k_1^4 k_2^3 + k_2^4 k_3^3 + k_3^4 k_0^3} = q_0 q_1 \frac{q_1 + q_0^2 + (1 + q_0 q_1^2) r_1}{1 + q_0^4 q_1^3 + (q_0^2 + q_1^4) r_1},$$

woraus sich k_0 und r_1 als rationale Funktionen der q_0, q_1, r_1, φ_1 bzw. der q_0, q_1, φ_2 berechnen lassen. Für die zyklischen Funktionen der q_0, q_1 bilden

$$(18) \quad \psi_3 = q_0 q_1^2, \quad \psi_4 = q_1 q_0^{-2} \quad \text{oder} \quad \chi_3 = \frac{\psi_3 - 1}{\psi_3 + 1}, \quad \chi_4 = \frac{\psi_4 - 1}{\psi_4 + 1} \text{(vgl. 16, 16a)}$$

eine Basis, da mit ihnen auch z. B. $q_0^5 = \psi_3 \psi_4^{-2}$ bekannt ist. Da T unter ihnen die Substitution $(\chi_4, \chi_3, -\chi_4, -\chi_3)$ hervorruft, so bilden weiter die Funktionen

$$(19) \qquad \chi_3^2, \ \chi_3 \chi_4$$

---

[8]) Für $n = 2$ sind dies die halbmetazyklischen Funktionen.

[9]) Vgl. E. Fischer, Zur Theorie der endlichen Abelschen Gruppen, Mathematische Annalen 77. Daselbst werden jedoch nur solche Gruppen von Cremona-Transformationen behandelt, die — was *hier* nicht notwendig der Fall ist — durch Einführung einer geeigneten Basis in eine Gruppe linearer Transformationen verwandelt werden können.

[10]) Der Anfang einer solchen Tabelle findet sich bei Wäisälä a. a. O. S. 31.

eine Basis für die halbmetazyklischen Funktionen und ebenso die Größen

$$(20) \qquad \bar{\varphi}_3 = \frac{\chi_3^2 - \chi_4^2}{2\chi_3\chi_4}, \qquad \bar{\varphi}_4 = \frac{\chi_3^2 + \chi_4^2}{2}$$

eine Basis für die vollmetazyklischen Funktionen der q_0, q_1. Genau so wie wir hier den Körper der rationalen Funktionen der Größen q_0, q_1 allmählich zu dem aus (20) hervorgehenden eingeschränkt haben, verläuft umgekehrt, ganz entsprechend der Zerlegung der linearen Gruppe in ihre Faktorgruppen, der Aufstieg von dem letztgenannten Körper zu dem ersten durch Adjunktion zweier Quadratwurzeln und einer fünften Wurzel:

$$(21) \qquad \chi_3\chi_4 = \frac{\bar{\varphi}_4}{\sqrt{\bar{\varphi}_3^2+1}}, \quad \begin{matrix}\chi_3^2\\ \chi_4^2\end{matrix} = \bar{\varphi}_4 \pm \bar{\varphi}_3 \cdot \chi_3\chi_4; \quad \chi_3 = \sqrt{\chi_3^2}; \quad q_0 = \sqrt[5]{\psi_3\psi_4^{-2}}.$$

Von den Basen (18), (19), (20) besitzt jedoch nur die letzte (20) Koeffizienten aus dem absoluten Rationalitätsbereich, während die Größen (18), (19) sich noch ändern, wenn man ε durch ε^g ersetzt: es geht χ_3 in χ_4, χ_4 in $-\chi_3$ über. (Vgl. das nach Formel (7), (8) Bemerkte.)[11]) Setzen wir aber

$$\eta_0 = \varepsilon - \varepsilon^{-1}, \qquad \eta_1 = \varepsilon^2 - \varepsilon^{-2},$$

---

[11]) Die Formeln (9a), (13a), (18), (19), (20) geben, von der Bezeichnung abgesehen, die urspünglichen Noetherschen Formeln für $p=5$ wieder, ebenso weiter unten die Formeln (9b), (10b) und (25b) in Fußnote 13 die für den Fall $p=7$. Der Verfasser erhielt nun, indem er der Basis (20) die rationale halbmetazyklische Funktion $(\varepsilon+\varepsilon^4-\varepsilon^2-\varepsilon^3)\chi_3\chi_4$ hinzufügte, eine halbmetazyklische Rationalbasis, und ebenso durch Adjunktion von α_3 (18a) eine zyklische Rationalbasis mit zunächst je sechs Basisfunktionen, die aber sodann mit Hilfe der zwischen ihnen bestehenden Gleichung leicht auf fünf reduziert werden konnten, womit für beide Zwischenkörper *rationale* Minimalbasen gewonnen waren. Fräulein Noether bemerkte sodann, daß diese Rechnung zu ersparen ist, indem man neben α_3 gleichzeitig die durch die Substitution $(\varepsilon, \varepsilon^2)$ in den η aus ihr hervorgehende Funktion α_4 einführt. In gleicher Weise hatte der Verfasser weiter unten im Falle $p=7$ immer nur *eine*, zu dem betreffenden Zwischenkörper gehörige Funktion eingeführt und sodann durch Rechnung die sämtlichen rationalen Minimalbasen bestimmt, während Fräulein Noether sodann bemerkte, daß man nur die drei Funktionen (23a) gleichzeitig einzuführen braucht, um ohne Rechnung zum gleichen Ziele zu gelangen. Die von ε freie Basis Q_3, Q_4 des Körpers der Unbestimmten selbst und die für die halbmetazyklischen Funktionen waren dann leicht auf dem gleichen Wege zu bestimmen. Daß die hier benutzten Funktionen, z. B. α_3 (18a), von ε frei sind, rührt davon her, daß die Größen η_0, η_1 unter $(\varepsilon, \varepsilon^2)$ die gleiche Substitution erfahren wie χ_3, χ_4. Eine lineare Kombination beider bleibt daher ungeändert. Im wesentlichen das gleiche Prinzip, nur noch auf die Anzahl p statt hier auf $\frac{p-1}{2}$ angewendet, liegt schon der Formel (3) zugrunde, in der die x_ν aus den Lagrangeschen Resolventen k_μ und den Größen $\varepsilon^{-g^{\mu}\cdot\nu}$ linear zusammengesetzt wurden, welche unter $(\varepsilon, \varepsilon^g)$ die gleiche zyklische Vertauschung erleiden.

wobei unter ε die (zu Beginn des § 1) fest gewählte 5-te Einheitswurzel verstanden ist, so leisten die Funktionen

$$(18a)\qquad \alpha_3 = \eta_0 \chi_3 + \eta_1 \chi_4, \quad \alpha_4 = \eta_1 \chi_3 - \eta_0 \chi_4$$

$$(19a)\qquad \alpha_3^2, \quad \alpha_3 \alpha_4$$

$$(20a)\qquad \varphi_3 = \frac{\alpha_3^2 - \alpha_4^2}{2\alpha_3 \alpha_4}, \quad \varphi_4 = \frac{\alpha_3^2 + \alpha_4^2}{2}$$

offenbar den gleichen Dienst wie die Funktionen (18), (19), (20), sind aber im Gegensatz zu diesen *sämtlich* frei von ε, da sich bereits α_3 und α_4 nicht ändern, wenn man ε durch ε^2 ersetzt. Die Funktionen (20a) unterscheiden sich übrigens nur unwesentlich von denen (20). Bezeichnen wir weiter

$$Q_0 = \frac{q_0 - 1}{q_0 + 1}, \quad Q_1 = \frac{q_1 - 1}{q_1 + 1}; \quad Q_3 = \eta_0 Q_0 + \eta_1 Q_1; \quad Q_4 = \eta_1 Q_0 - \eta_0 Q_1,$$

so bilden Q_3, Q_4 eine Basis für diejenigen rationalen Funktionen der q_0, q_1, die als Funktionen der x_ν Koeffizienten aus dem absoluten Rationalitätsbereich haben, da einerseits auch q_0, q_1 mit Hilfe von ε rational durch Q_3, Q_4 ausdrückbar sind, letztere aber sich nicht ändern, wenn man ε durch ε^2 ersetzt. Mithin bilden schließlich

$$(22,1)\qquad \varphi_0, \varphi_1, \varphi_2, Q_3, Q_4$$

$$(22,2)\qquad \varphi_0, \varphi_1, \varphi_2, \alpha_3, \alpha_4$$

$$(22,3)\qquad \varphi_0, \varphi_1, \varphi_2, \alpha_3^2, \alpha_3 \alpha_4$$

$$(22,4)\qquad \varphi_0, \varphi_1, \varphi_2, \varphi_3, \varphi_4$$

je eine Minimalbasis für die rationalen Funktionen der x_ν überhaupt, bzw. für die zyklischen, halbmetazyklischen und vollmetazyklischen Funktionen der x_ν, und sie besitzen sämtlich Koeffizienten aus dem absoluten Rationalitätsbereich. Der Aufstieg vom Körper (22, 4) zu (22, 1) erfolgt jeweils durch Änderung nur der beiden letzten Basisfunktionen, er entspricht genau dem durch die Gleichungen (21) dargestellten und ist aus diesen fast unmittelbar abzulesen bzw. zu übertragen.

Im Falle $p = 7$ wählen wir $g = 3$ und finden, gleichfalls durch Tabulierung,

$$(9b)\qquad \varphi_1 = \frac{k_0^2 k_3 + \dots + k_5^2 k_4 + 2(k_0 k_2 k_4 + k_1 k_3 k_5)}{k_0 k_3 + k_1 k_4 + k_2 k_5} \quad \text{(vgl. § 3, Fußnote 15)},$$

$$(10b)\qquad \varphi_2 = \frac{k_0^3 k_1 k_2^3 k_4 + \dots + k_5^3 k_0 k_1^2 k_3}{k_0^3 k_4^2 k_5^2 + \dots + k_5^3 k_3^2 k_4^2}, \quad \varphi_3 = \frac{k_0^2 k_1^2 k_2 k_3 k_5 + \dots + k_5^2 k_0^2 k_1 k_2 k_4}{k_0^3 k_1 k_2^2 k_4 + \dots + k_5^3 k_0 k_1^2 k_3}.$$

Aus (10b) erhalten wir bei Kürzung durch k_0^7 zwei lineare Gleichungen für r_1, r_2, aus denen letztere „berechnet“ werden können.

Für die zyklischen Funktionen der q_0, q_1, q_2 bilden

$$(23)\quad \psi_4 = q_0 q_1^2,\ \psi_5 = q_1 q_2^2,\ \psi_6 = q_2 q_0^{-2} \quad \text{oder} \quad \chi_e = \frac{\psi_e - 1}{\psi_e + 1} \ldots (e = 4, 5, 6),$$

(vgl. (16), (16a)) eine Basis. Die χ_e werden durch T in der Reihenfolge $(\chi_6, \chi_5, \chi_4, -\chi_6, -\chi_5, -\chi_4)$ vertauscht, genügen also einer rationalen zyklischen Gleichung 6. Grades, in der nur die quadratischen Glieder auftreten. Daher [12]) dulden

$$(24)\qquad \zeta_4 = \chi_5 \chi_6, \qquad \zeta_5 = -\chi_6 \chi_4, \qquad \zeta_6 = \chi_4 \chi_5 \qquad (\text{vgl. } 17)$$

die Vertauschung T^3 und bilden eine Basis der gegenüber S und T^3 invarianten Funktionen der q_0, q_1, q_2. Sie werden durch T nur zyklisch vertauscht und somit bilden

$$(25)\quad \begin{cases} \bar{\varphi}_4 = \dfrac{\zeta_4 + \zeta_5 + \zeta_6}{3}, \qquad \bar{\varphi}_5 = -\dfrac{3}{2} \dfrac{\bar{\zeta}_4 \bar{\zeta}_5 \bar{\zeta}_6}{\bar{\zeta}_4 \bar{\zeta}_5 + \bar{\zeta}_5 \bar{\zeta}_6 + \bar{\zeta}_6 \bar{\zeta}_4}, \\[2ex] \qquad \bar{\varphi}_6 = -\dfrac{1}{6} \dfrac{(\zeta_4 - \zeta_5)(\zeta_5 - \xi_6)(\zeta_6 - \zeta_4)}{\bar{\zeta}_4 \bar{\zeta}_5 + \bar{\zeta}_5 \bar{\zeta}_6 + \bar{\zeta}_6 \bar{\zeta}_4}, \end{cases}$$

wobei $\zeta_e - \bar{\varphi}_4 = \xi_e$ gesetzt ist, eine Basis für die metazyklischen Funktionen der q_0, q_1, q_2 [13]). Den Formeln (21) entsprechen, wie unten noch näher ausgeführt wird, hier die folgenden, in denen $\bar{\xi}_4$ nur zur Vereinfachung der Schreibweise herausgegriffen ist, ohne jedoch eine bevorzugte Rolle zu spielen:

$$(26)\quad \begin{cases} \bar{\xi}_4 = A + \dfrac{\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2}{A}, \quad \text{wobei } A = \sqrt[3]{(\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2)(\bar{\varphi}_5 + \bar{\varphi}_6 \sqrt{-3})}, \\[2ex] \bar{\xi}_{5,6} = -\dfrac{1}{2}\bar{\xi}_4 \mp 3\bar{\varphi}_6 \dfrac{\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2}{\bar{\zeta}_4^2 - (\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2)}, \qquad \zeta_e = \bar{\varphi}_4 + \bar{\xi}_e; \\[2ex] \chi_4 = \sqrt{-\dfrac{\zeta_5 \zeta_6}{\zeta_4}}; \quad q_0 = \sqrt[7]{\dfrac{\psi_5^2}{\psi_4 \psi_6^4}}. \end{cases}$$

Daß hier in der Größe A die $\sqrt{-3}$ auftritt, darf nicht verwundern. Die „Entfernung des Faktors 3" aus der linearen Gruppe erfolgt rein durch Adjunktion einer Wurzel der Abelschen Gleichung *dritten* Grades, der die $\bar{\xi}_e$ genügen. Bei der Darstellung dieser Wurzel durch *Radikale*

---

[12]) Zu den hier folgenden Formeln, die von den Noetherschen etwas abweichen (siehe Einleitung) vgl. Breuer, Zyklische Gleichungen 6. Grades und Minimalbasis, Math. Annalen 86, S. 108ff.

[13]) In den Noetherschen Formeln traten an Stelle von (25) die folgenden Funktionen auf:

$$(25\text{b})\qquad \varphi_4' = \frac{\xi_4^2 \xi_5 - \xi_5^2 \xi_6 - \xi_6^2 \xi_4}{\xi_4 \xi_5 \xi_6}, \quad \varphi_5' = \xi_4^2 + \xi_5^2 + \xi_6^2, \quad \varphi_6' = \frac{\varrho}{\xi_4 \xi_5 \xi_6},$$

wobei

$$\varrho = \chi_4 - \chi_5 + \chi_6, \quad \xi_4 = \chi_4 - \frac{\varrho}{3}, \quad \xi_5 = \chi_5 + \frac{\varrho}{3}, \quad \xi_6 = \chi_6 - \frac{\varrho}{3}.$$

tritt aber die $\sqrt{-3}$ als akzessorische Irrationalität auf. Die zyklische Gleichung für ξ, die zu den Formeln (26) führt, hat die Form:

$$\xi^3 - 3(\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2)\xi - 2\bar{\varphi}_5(\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2) = 0 \tag{27}$$ [12].

Nun gilt für jede kubische Gleichung $f(y) \equiv y^3 + a_1 y^2 + a_2 y + a_3 = 0$, daß $y_{2,3} = -\frac{1}{2}\left(a_1 + y_1 \pm \frac{\Delta}{f'(y_1)}\right)$, wobei unter Δ das Differenzenprodukt $(y_1 - y_2)(y_2 - y_3)(y_3 - y_1)$ verstanden ist und das obere Vorzeichen sich auf y_2 bezieht. Durch Anwendung auf die Gleichung (27), für welche das Differenzenprodukt $\Delta \equiv (\xi_4 - \xi_5)(\xi_5 - \xi_6)(\xi_6 - \xi_4) = +18\bar{\varphi}_6(\bar{\varphi}_5^2 + 3\bar{\varphi}_6^2)$ ist, wie sich aus der Bedeutung von $\bar{\varphi}_6$ in (25) ergibt, wird die mittlere Formel (26) erhalten, in der sich also das obere Vorzeichen auf ξ_5 bezieht. Bezeichnen wir mit G_{42}, G_{21}, G_{14}, G_7 die durch S und T, bzw. S und T^2, bzw. S und T^3, bzw. S erzeugte Permutationsgruppe, so gehören also die Größen $\bar{\varphi}_e$ (25) und φ'_e (25b) zu G_{42}, die Größen ζ_e, ξ_e (24), (25) zu G_{14}, ϱ (25b) zu G_{21}, endlich die Größen χ_e (23) und ξ_e (25b) zu G_7. Zwei Kompositionsreihen von G_{42} werden durch G_{42}, G_{21}, G_7, E und G_{42}, G_{14}, G'_7, E dargestellt. Die zugehörigen Faktorgruppen sind isomorph zu zyklischen Permutationsgruppen 2., 3., 7., bzw. 3., 2., 7. Ordnung. Die zweite Kompositionsreihe liegt offenbar *unserer* Auflösung zugrunde, die Gleichungen, aus denen in (26) die Größen ξ_e und χ_e gewonnen werden, besitzen zyklische Gruppen vom 3. bzw. 2. Grade. In den nicht ganz so durchsichtigen Noetherschen Formeln (25b) wird die erste Kompositionsreihe benutzt, die unschwer aufzustellenden Gleichungen, denen ϱ bzw. nach dessen Adjunktion die Größen ξ_e genügen, besitzen zyklische Gruppen vom 2. bzw. 3. Grade. Die Gruppen der einzuführenden Hilfsgleichungen sind also isomorph zu den Faktorgruppen der jeweils benutzten Kompositionsreihe, wie dies nicht anders zu erwarten war.

In den Formeln (25b, Fußnote 13) tritt für die aus S und T^2 erzeugte halbmetazyklische Zwischengruppe eine Basis nicht explizit auf. Der Vollständigkeit halber und mit Rücksicht auf das Folgende bemerken wir, daß sie sich in einfachster Weise bilden läßt als Minimalbasis der die Vertauschung $(\chi_4, -\chi_5, \chi_6)$ duldenden Funktionen dieser Größen, d. h. analog zu Formel (25) etwa in der Form

$$\left\{\begin{aligned} &\omega_4 = \frac{1}{3}\varrho = \frac{\chi_4 - \chi_5 + \chi_6}{3}, \qquad \omega_5 = -\frac{3}{2}\,\frac{\xi_4\xi_5\xi_6}{\xi_4\xi_5 + \xi_5\xi_6 - \xi_6\xi_4}, \\ &\qquad \omega_6 = -\frac{1}{6}\,\frac{(\chi_4 + \chi_5)(\chi_5 + \chi_6)(\chi_6 - \chi_4)}{\xi_4\xi_5 + \xi_5\xi_6 - \xi_6\xi_4}. \end{aligned}\right. \tag{28}$$

Genau wie vorhin im Falle $p = 5$, so besitzt auch hier diese Basis (28) so wenig wie die Basen (23), (24) Koeffizienten aus dem absoluten Ra-

tionalitätsbereich, und nur die vollmetazyklische Basis (25) selbst — ebenso natürlich die in Fußnote 13 auftretende vollmetazyklische Basis (25 b) — hat rationale Koeffizienten. Setzen wir aber wieder

$$\eta_0 = \varepsilon - \varepsilon^{-1}, \qquad \eta_1 = \varepsilon^3 - \varepsilon^{-3}, \qquad \eta_2 = \varepsilon^2 - \varepsilon^{-2},$$

so leisten die Funktionen

$$(23\text{a}) \qquad \begin{cases} \alpha_4 = \eta_0\chi_4 + \eta_1\chi_5 + \eta_2\chi_6, \qquad \alpha_5 = \eta_1\chi_4 + \eta_2\chi_5 - \eta_0\chi_6, \\ \alpha_6 = \eta_2\chi_4 - \eta_0\chi_5 - \eta_1\chi_6, \end{cases}$$ [14]

$$(24\text{a}) \qquad \beta_4 = \alpha_5\alpha_6, \qquad \beta_5 = -\alpha_6\alpha_4, \qquad \beta_6 = \alpha_4\alpha_5$$

offenbar denselben Dienst wie die Funktionen (23), (24), sind aber im Gegensatz zu jenen frei von ε. Die Funktionen

$$(25\text{a}) \qquad \varphi_4, \varphi_5, \varphi_6,$$

die sich aus den β_e ebenso zusammensetzen wie die Größen (25) aus den ζ_e, können an deren Stelle treten, unterscheiden sich übrigens nur unwesentlich von ihnen. Bilden wir ferner Größen

$$(28\text{a}) \qquad \gamma_4, \gamma_5, \gamma_6,$$

die sich aus den α_e (23a) ebenso zusammensetzen wie die ω_e (28) aus den χ_e (23) [wobei zu beachten ist, daß nach Fußnote 13 $\xi_4 = \chi_4 - \frac{\varrho}{3}$ usw.], so erhält man eine gleichfalls von ε freie halbmetazyklische Basis.

Bezeichnen wir nun noch, wie vorher im Falle $p = 5$,

$$Q_0 = \frac{q_0 - 1}{q_0 + 1}, \qquad Q_1 = \frac{q_1 - 1}{q_1 + 1}, \qquad Q_2 = \frac{q_2 - 1}{q_2 + 1};$$

$$Q_4 = \eta_0 Q_0 + \eta_1 Q_1 + \eta_2 Q_2, \qquad Q_5 = \eta_1 Q_0 + \eta_2 Q_1 - \eta_0 Q_2,$$
$$Q_6 = \eta_2 Q_0 - \eta_0 Q_1 - \eta_1 Q_2,$$

so bilden wiederum Q_4, Q_5, Q_6 eine Basis für diejenigen rationalen Funktionen der q_0, q_1, q_2, die als Funktionen der x_ν Koeffizienten aus dem absoluten Rationalitätsbereich besitzen. Mithin bilden

$$(29, 1) \qquad \varphi_0, \varphi_1, \varphi_2, \varphi_3, Q_4, Q_5, Q_6;$$

$$(29, 2) \qquad \varphi_0, \varphi_1, \varphi_2, \varphi_3, \alpha_4, \alpha_5, \alpha_6;$$

$$(29, 3) \quad \varphi_0, \varphi_1, \varphi_2, \varphi_3, \beta_4, \beta_5, \beta_6; \qquad (29, 4) \quad \varphi_0, \varphi_1, \varphi_2, \varphi_3, \gamma_4, \gamma_5, \gamma_6;$$

$$(29, 5) \qquad \varphi_0, \varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \varphi_6$$

je eine Minimalbasis für die rationalen Funktionen der x_ν überhaupt bzw. für die Invarianten gegenüber S, S und T^3, S und T^2, S und T. Der

[14]) Die Determinante dieser drei Gleichungen ist, ebenso wie oben die der Gleichungen (18a), von Null verschieden.

Aufstieg von den letzteren, nämlich dem vollmetazyklischen Körper (29, 5) über den halbmetazyklischen (29, 4) oder den gegenüber S und T^3 invarianten Körper (29, 3) zum zyklischen (29, 2) und weiter zum „rationalen" Körper (29, 1), vollzieht sich durch Adjunktionen, d. h. durch Auflösung von Hilfsgleichungen, die nach dem früher Gesagten leicht aufgestellt und daher hier übergangen werden können.

§ 3.

Das im vorstehenden behandelte Problem war, wie nochmals hervorgehoben sei, ein solches der Körper gewisser Funktionen von p Unbestimmten. Wir konnten die Gewinnung einer Basis für die *verschiedenen* bei einer und derselben Anzahl p der Unbestimmten zu untersuchenden Körper vereinfachen, indem wir zeigten, daß die $n+1$ *ersten* Basisfunktionen für alle zu untersuchenden Körper — nämlich die der rationalen und der metazyklischen Funktionen und alle Zwischenkörper — gemeinsam auftreten, so daß nur die restlichen n Basisfunktionen mit der Natur des Körpers sich änderten und diese bestimmten. Ein ganz anderes Problem — für dessen Lösung freilich die Existenz der Basis eine hinreichende Voraussetzung bildet[1]) — liegt aber vor, wenn die Koeffizienten und Wurzeln der in irgendeinem Bereich metazyklischen Gleichungen durch p Parameter dargestellt werden sollen. Denn dann sucht man nicht eine Basis für den Körper der metazyklischen Funktionen überhaupt nebst dem möglichst bequemen Aufstieg zum *Körper* der Unbestimmten selbst, sondern man sucht für alle *speziellen* Größen x_ν, deren metazyklische Funktionen dem gegebenen Rationalitätsbereich angehören, eine Darstellung *spezieller* metazyklischer — nämlich der symmetrischen — Funktionen durch geeignete Basisfunktionen, und sucht weiter den Aufstieg von letzteren zu den Größen x_ν selbst, nicht nur zu deren Körper. Hierzu ist der von Herrn Wäisälä[2]) eingeschlagene Weg der geeignetere. Herr Wäisälä drückt zunächst nach dem Vorgang von Euler die Koeffizienten (ebenso auch die Wurzeln) auflösbarer Gleichungen 5. Grades durch die k_μ aus und führt bei ersteren für gewisse zyklische Verbindungen der k_μ, die also Wurzeln rationaler zyklischer biquadratischer Gleichungen sein müssen, die Webersche bzw. Abelsche Parameterdarstellung solcher Wurzeln ein. Dadurch ergibt sich auf verhältnismäßig einfachem Wege eine Parameterdarstellung für die Koeffizienten der Gleichung 5. Grades, zu der freilich auch mehrere, von Herrn W. gleichfalls ausgeführte Ergänzungsdarstellungen erforderlich werden. Nach dem Vorstehenden kann es nicht überraschen — wenngleich Herr W. diese Tatsache nicht erwähnt hat —, daß auch die bei Herrn W. zuletzt auftretenden Parameter, obwohl auf ganz anderem Wege gewonnen, selbst wieder metazyklische Funktionen der Wurzeln sind und

somit eine Minimalbasis bilden (s. u.). Umgekehrt kann man nicht erwarten, daß unsere in § 2 gewonnenen Basisfunktionen zur Koeffizientendarstellung besonders geeignet seien. Betrachtet man, was ohne wesentliche Einschränkung der Allgemeinheit möglich ist, nur Gleichungen der Form

$$x^p + a_2 x^{p-2} + a_3 x^{p-3} + \ldots + a_{p-1} x + a_p = 0,$$

so stehen freilich die Funktionen φ_0 (2) und φ_1 (9a), (9b) in einfacher Beziehung zu den Koeffizienten, da $\varphi_0 = 0$, $\varphi_1 = \frac{a_3}{a_2}$ [15]). Aber schon die Berechnung des Koeffizienten a_2 für $p = 5$ wird vollkommen unübersichtlich. Um aber auch nur zu der *Basis* (22, 4) alle für *spezielle* Werte x_ν erforderlichen Ergänzungsdarstellungen anzugeben, müßte man vorerst die Koeffizienten durch die Basis (22, 4) ausdrücken. Denn letztere wird nicht nur für solche Werte unbrauchbar, in denen etwa einzelne Basisfunktionen verschwinden oder unendlich werden — z. B. also für $a_2 = 0$ wegen $\varphi_1 = \infty$, desgleichen für $a_3 = 0$ wegen $k_0 = \infty$ —, sondern auch für solche Werte x_ν, die etwaige in den Ausdrücken der Koeffizienten auftretende Nenner zum Verschwinden bringen. Wir verzichten daher, namentlich auch im Hinblick auf die vorliegende Wäisäläsche Darstellung, auf Durchführung der Diskussion unserer Basis (22, 4) und wollen nur beispielsweise die Ergänzungsdarstellungen angeben, die für den Parameter 1. Grades φ_1 nötig werden, wobei wir der Einfachheit halber alle k_μ als von 0 verschieden voraussetzen. Um hier nur eine einzige Ausnahme betrachten zu müssen, wählen wir $\varphi_1' = \frac{a_2^3}{a_5}$ $\left(\text{statt } \varphi_1 = \frac{a_3}{a_2}\right)$, brauchen — da uns nur irreduzible Gleichungen interessieren — nur

$$1. \quad a_3 = 0$$

in Betracht zu ziehen und setzen in diesem Falle $\varphi_1'' = \frac{a_5}{a_2^2}$. Alsdann brauchen wir keine Basisfunktion φ_2 einzuführen, da die Bedingung $a_3 = 0$, d. h. $k_0^2 k_3 + \ldots + k_3^2 k_2 = 0$, allein hinreicht, um r_1 durch q_0, q_1 auszudrücken in der Form $r_1 = -\frac{q_0^2 + q_1}{1 + q_0 q_1^2}$. Es bilden also in diesem Falle $\varphi_0, \varphi_1'', \varphi_3, \varphi_4$ die Minimalbasis. Ist aber

$$2. \quad a_2 = a_3 = 0,$$

[15]) Zu (9a) vgl. Wäisälä a. a. O. S. 32 oder Breuer a. a. O. Fußnote 3, S. 167. Für (9b) ergibt eine einfache Rechnung die Richtigkeit unserer Behauptung. Wir hätten natürlich in (9a), (9b) unmittelbar $\varphi_1 = \frac{a_3}{a_2}$ oder gleich sonst irgendeiner symmetrischen Funktion 1. Grades der x_ν setzen können, ohne diese erst durch die k_μ auszudrücken.

so setzen wir $\varphi_1''' = \frac{a_3}{a_4}$, wonach wiederum die k_μ durch φ_1''', q_0, q_1, r ausgedrückt werden können. Aus $a_2 = 0$, d. h. $k_0 k_2 + k_1 k_3 = 0$ folgt $q_0 + q_1 r_1^2 = 0$. Ersetzt man hierin r_1 durch den soeben gewonnenen Ausdruck, so erhält man aus (18) leicht die Beziehung

$$\psi_4 (1 + \psi_3)^2 + \psi_3 (1 + \psi_4)^2 = 0, \tag{30}$$

die gestattet, an Stelle von ψ_3, ψ_4 einen einzigen Parameter

$$\psi_2 = \frac{\psi_3 + 1}{\psi_4 + 1}$$

zu setzen; denn führt man diesen in (30) ein, so wird erhalten

$$\psi_3 = \frac{\psi_2 - 1}{\psi_2 + 1}\,\psi_2, \qquad \psi_4 = \frac{\psi_2 - 1}{\psi_2 + 1} \cdot \frac{-1}{\psi_2}.$$

Setzt man noch

$$2\,\omega_2 = \psi_2 - \frac{1}{\psi_2}, \qquad 2\,\varphi_2 = \omega_2 - \frac{1}{\omega_2},$$

so erkennt man leicht, daß die Funktionen

$$\varphi_0,\ \varphi_1''',\ \psi_2; \qquad \varphi_0,\ \varphi_1''',\ \omega_2; \qquad \varphi_0,\ \varphi_1''',\ \varphi_2$$

die entsprechenden früheren ersetzen, ebenso die Gleichungen

$$\omega_2 = \varphi_2 + \sqrt{\varphi_2^2 + 1}, \qquad \psi_2 = \omega_2 + \sqrt{\omega_2^2 + 1}, \qquad q_0 = \sqrt[5]{\psi_2^3\,\frac{\psi_2 + 1}{\psi_2 - 1}}$$

die Gleichung (21).

3. Der Fall $a_2 = a_3 = a_4 = 0$ führt auf die binomischen Gleichungen und bedarf keiner besonderen Behandlung.

Wir merken hier noch die Bedeutung der Basisfunktionen einer der bei Herrn Wäisälä implizite auftretenden Minimalbasen an. In dem ersten Hauptfall (a. a. O. S. 34, Formel (79)) sind dies neben unserem φ_0 (2), das dort gleich Null angenommen ist, die Parameter p, q, s, t, deren Bedeutung sich durch eine einfache Rechnung wie folgt ergibt, wenn wir hier mit Herrn Wäisälä speziell $v_\mu = k_\mu^2 k_{\mu+3}$ setzen:

$$(31,4)\quad \left\{\begin{aligned} p &= \frac{(v_0 + v_2) + (v_1 + v_3)}{k_0 k_2 - k_1 k_3} \cdot \frac{(v_0 - v_2)^2 + (v_1 - v_3)^2}{4\,(v_0 - v_2)(v_1 - v_3)} \\ &= -\frac{a_3}{20} \cdot \frac{(v_0 - v_2)^2 + (v_1 - v_3)^2}{(k_0 k_2 - k_1 k_3)(v_0 - v_2)(v_1 - v_3)}, \\ q &= \frac{(v_0 + v_2) - (v_1 + v_3)}{2\,(k_0 k_2 - k_1 k_3)}, \qquad s = \frac{(v_0 - v_2)^2 - (v_1 - v_3)^2}{2\,(v_0 - v_2)(v_1 - v_3)}, \\ t &= \frac{(v_0 - v_2)^2 + (v_1 - v_3)^2}{2\,(k_0 k_2 - k_1 k_3)^2}. \end{aligned}\right.$$

Die Formeln (31, 4) können also an Stelle von (22, 4) treten, wobei in ersteren φ_0 hinzuzufügen ist. Der Aufstieg zum Körper der halbmetazyklischen Funktionen erfolgt durch Adjunktion der Quadratwurzel $\sqrt{a} = \sqrt{1+s^2} = \frac{(v_0-v_2)^2+(v_1-v_2)^2}{2(v_0-v_2)(v_1-v_2)}$ (vgl. (21)), so daß zunächst die Größen

$$\varphi_0,\ p,\ q,\ s,\ t,\ \sqrt{1+s^2}$$

eine halbmetazyklische Basis bilden, die aber noch sechs Basisfunktionen enthält. Setzt man aber zur Abkürzung

$$(32)\qquad u = s + \sqrt{1+s^2} = \frac{v_0-v_2}{v_1-v_2},$$

so wird $2\sqrt{1+s^2} = u + \frac{1}{u}$, $2s = u - \frac{1}{u}$, d. h. die Größen s und $\sqrt{1+s^2}$ sind beide durch u allein rational darstellbar, und die Größen

$$(31,3)\qquad \varphi_0,\ p,\ q,\ t,\ u$$

bilden eine halbmetazyklische *Minimal*basis, deren Koeffizienten aber, ganz wie dies für (19) galt, im Gegensatz zu denen von (31, 4) nicht mehr von ε frei sind, wie sich leicht ergibt. Der Aufstieg zum Körper der zyklischen Funktionen wird dann weiter durch Ausziehen z. B. der Quadratwurzel

$$(33)\qquad w = \sqrt{t(a + s\sqrt{a})} = \frac{(v_0-v_2)^2+(v_1-v_2)^2}{2(v_1-v_2)(k_0k_2-k_1k_3)}$$

vollzogen, so daß

$$(31,2)\qquad \varphi_0,\ p,\ q,\ u,\ w$$

eine zyklische Minimalbasis bilden. Denn da mit u (32) auch s und $\sqrt{a}$ bekannt sind, ist mit u, w auch t bekannt nach (33)[16]. Durch Adjunktion einer fünften Wurzel, deren Radikand sich rational aus den Größen (31, 2) zusammensetzt (vgl. a. a. O. bei Herrn Wäisälä, Formel (81)), wird sodann z. B. k_0 selbst bekannt und damit einerseits der Aufstieg zum Körper der x_ν vollzogen, andererseits eine sehr einfache Darstellung der Größen x_ν selbst gewonnen.

Nach dem (zu Eingang dieses Paragraphen kurz beschriebenen) Verfahren des Herrn Wäisälä könnte man nun theoretisch auch die Gleichung 7. Grades behandeln — die Berechnung der Koeffizienten aus *unseren* Basisfunktionen φ_ν ist natürlich so wenig ratsam wie bei der Gleichung 5. Grades —, da die Wurzeln zyklischer Gleichungen 6. Grades als Funk-

[16]) Die vorstehenden Angaben wurden mit Rücksicht darauf so ausführlich gehalten, daß Herr Wäisälä, entsprechend *seiner* Problemstellung, die hier behandelten Fragen der Invariantenkörper und deren Minimalbasen überhaupt nicht berührt.

tionen von sechs Parametern bekannt sind [12]) und es auch keine Schwierigkeiten bereitet, die Koeffizienten der metazyklischen Gleichungen 7. Grades durch die k_μ auszudrücken. Während jedoch die Ausdrücke der Potenzsummen durch die k_μ — aus denen man zweckmäßig die der Koeffizienten selbst erst berechnet — bei der Gleichung 5. Grades nur aus einem bzw. drei „Leitgliedern" v_0 (6) bestehen, setzen sich die Potenzsummen $s_2, s_3, \ldots, s_7$ für die Gleichung 7. Grades, wie eine Aufstellung der Formeln ergibt, aus bzw. 1, 2, 4, 6, 9, 18 Leitgliedern zusammen. Es würden daher schon die rein rechnerischen Schwierigkeiten die vollständige Durchführung äußerst mühsam gestalten.

(Eingegangen am 1. 2. 1924.)

Über eine Klasse meromorpher Funktionen.

Von

Rolf Nevanlinna in Helsingfors.

Auf den folgenden Seiten wird ein Kriterium dafür gegeben, daß eine analytische Funktion, die innerhalb eines zusammenhängenden Gebietes G eindeutig und meromorph ist, sich als Quotient von zwei *beschränkten* Funktionen darstellen läßt. In dem besonderen Fall, wo G ein Kreis ist, können unsere Ergebnisse in folgendem Satze zusammengefaßt werden:

Sei $f(x)$ eine analytische Funktion der komplexen Variablen $x = re^{i\varphi}$, die für $|x| < R$ eindeutig und meromorph ist und für $x = 0$ einen endlichen Wert annimmt. Es bezeichne ferner $n(t)$ die Anzahl der Pole von $f(x)$ für $|x| < t$, und $\overset{+}{\log}|f|$ die Zahl $\log|f|$ oder Null, je nachdem $|f| \geqq 1$ oder $|f| < 1$.

Unter diesen Voraussetzungen ist die Summe

$$T(r) = \frac{1}{2\pi}\int_0^{2\pi} \overset{+}{\log}|f(re^{i\varphi})|\,d\varphi + \int_0^r \frac{n(t)}{t}\,dt$$

eine wachsende, konvexe Funktion von $\log r$.

Notwendig und hinreichend, damit die Funkton $f(x)$ als Quotient von zwei innerhalb des Kreises $|x| < R$ beschränkten Funktionen dargestellt werden kann, ist, daß $T(r)$ für $r < R$ beschränkt ist.

Wenn $f(x)$ insbesondere für $|x| < R$ *regulär* ist, so verschwindet $n(t)$ identisch, und es wird also

$$T(r) = \frac{1}{2\pi}\int_0^{2\pi} \overset{+}{\log}|f(re^{i\varphi})|\,d\varphi.$$

In diesem speziellen Fall wurde der obige Satz früher in einer Arbeit von F. Nevanlinna und dem Verfasser bewiesen[1]) [2]).

Unter Anwendung eines bekannten Fatouschen Satzes über die Existenz von Randwerten beschränkter Funktionen folgt aus unserem Satze u. a., daß eine für $|x| < R$ meromorphe Funktion, für welche der Ausdruck $T(r)$ für $r < R$ beschränkt ist, bei radialer Annäherung an die Punkte $|x| = R$ fast überall wohlbestimmte Randwerte besitzt. Im Falle einer für $|x| < R$ regulären Funktion wurde dieser Satz zuerst in der in der Fußnote [1]) zitierten Arbeit bewiesen; unabhängig hiervon hat ihn auch Ostrowski[3]) gefunden.

1. Es sei $f(x)$ eine analytische Funktion der komplexen Variablen $x = re^{i\varphi}$, die innerhalb eines zusammenhängenden Gebietes G^* eindeutig und meromorph ist. Sei ferner G ein von einer endlichen Anzahl analytischer Kurvenstücke Γ begrenztes, zusammenhängendes Gebiet, dessen Rand vollständig innerhalb G^* verläuft; ein solches Gebiet werden wir kurz ein *inneres* Gebiet von G^* nennen. Wenn x ein innerer Punkt von G ist, so läßt sich $\log|f(x)|$ durch die Formel

$$(1)\qquad \log|f(x)| = \frac{1}{2\pi}\int_{\Gamma} \log|f(\xi)|\frac{\partial g(\xi, x)}{\partial n}\,ds - \sum_{G} g(x, a_\mu) + \sum_{G} g(x, b_\nu)$$

berechnen, wo $g(x, z)$ die Greensche Funktion des Gebietes G mit ihrem Pol in dem Punkt $x = z$ bezeichnet, und die Summen rechts über sämtliche innerhalb G gelegenen Nullstellen a_μ und Pole b_ν der Funktion $f(x)$ zu erstrecken sind[4]). Um diese Formel auf eine für die folgende Unter-

---

[1]) F. und R. Nevanlinna, Über die Eigenschaften analytischer Funktionen in der Umgebung einer singulären Stelle oder Linie, Acta Soc. Sc. Fennicae, **50** (1922), N. 5.

[2]) In einer neuerdings erschienenen Arbeit (Über die Randwerte einer analytischen Funktion, Math. Zeitschrift, **18**, Heft 1/2, (1923), S. 87–95) hat F. Riesz folgenden Zerlegungssatz bewiesen: Wenn $f(x)$ eine innerhalb des Einheitskreises reguläre, analytische Funktion ist, wofür das Integral $T(r)$ für $r \to 1$ beschränkt ist, so kann sie in der Form $f(x) = g(x)\,h(x)$ geschrieben werden, wo $h(x)$ für $|x| < 1$ beschränkt, und $g(x)$ eine im Einheitskreise nirgends verschwindende, reguläre Funktion ist, für welche der Mittelwert $T(r)$ ebenfalls beschränkt ist. Dieses Ergebnis ist offenbar in unserem oben angegebenen Satze (in der spezielleren Fassung, in welcher er schon in der in der Fußnote [1]) zitierten Arbeit bewiesen wurde) enthalten; nach diesem Satze ist nämlich $f = \frac{\varphi}{\psi}$, wo φ und ψ im Einheitskreise *beschränkt* sind, und $\psi \neq 0$ ist.

[3]) A. Ostrowski, Über die Bedeutung der Jensenschen Formel für einige Fragen der komplexen Funktionentheorie, Acta litt. ac scient. regiae univ. hung. Francisco-Josephinae **1** (1923), f. 2, S. 1–8.

[4]) Vgl. die in der Fußnote [1]) zitierte Arbeit von F. und R. Nevanlinna, insbesondere S. 7.

suchung zweckmäßige Form zu bringen, führen wir folgende Bezeichnungen ein:

$$(2)\qquad U_G(x,f)=\frac{1}{2\pi}\int_{\Gamma}\overset{+}{\log}|f(\xi)|\frac{\partial g(\xi,x)}{\partial n}ds,\qquad V_G(x,f)=\sum_{G} g(x,b_\nu).$$

Setzt man noch

$$(3)\qquad T_G(x,f)=U_G(x,f)+V_G(x,f),$$

so läßt sich die Formel (1) in der einfachen Form

$$(1)'\qquad \log|f(x)|=T_G(x,f)-T_G\left(x,\frac{1}{f}\right)$$

schreiben. Man bemerke, daß der Ausdruck $T_G(x,f)$ nach seiner Definition eine innerhalb G *nichtnegative*, harmonische Funktion ist, die auf Γ die Randwerte $\overset{+}{\log}|f(x)|$ annimmt und in den innerhalb G gelegenen Polen der Funktion $f(x)$ positive, logarithmische Stellen besitzt.

2. Der Ausdruck $T_G(x,f)$ besitzt folgende fundamentale Eigenschaft:

Satz I. *Wenn G_1 und G_2 innere Gebiete von G^* sind, und G_1 als Teilgebiet in G_2 enthalten ist, so gilt für jedes x innerhalb G_1*

$$T_{G_1}(x,f)\leqq T_{G_2}(x,f).$$

Zum Beweise nehmen wir einen beliebigen Punkt x_0 innerhalb G_1. Nach $(1)'$ ist in jedem Punkt x innerhalb G_2

$$\log|f(x)|=T_{G_2}(x,f)-T_{G_2}\left(x,\frac{1}{f}\right)\leqq T_{G_2}(x,f),$$

und, da die rechte Seite *nichtnegativ* ist, auch

$$\overset{+}{\log}|f(x)|\leqq T_{G_2}(x,f).$$

Diese Ungleichung besteht insbesondere in jedem Randpunkt von G_1. Multipliziert man beide Seiten mit $\frac{\partial g_1(x,x_0)}{\partial n}$, wo $g_1(x,x_0)$ die Greensche Funktion von G_1 ist, so folgt durch Integration längs der Randkurve Γ_1 von G_1, daß

$$(4)\qquad U_{G_1}(x_0,f)\leqq\frac{1}{2\pi}\int_{\Gamma_1}T_{G_2}(x,f)\frac{\partial g_1(x,x_0)}{\partial n}ds.$$

Zur Berechnung des letzten Integrals bemerke man, daß die Funktion $T_{G_2}(x,f)-V_{G_1}(x,f)$ innerhalb G_1 harmonisch und regulär ist und auf der Randkurve Γ_1 die Randwerte $T_{G_2}(x,f)$ annimmt. Daher wird:

$$\frac{1}{2\pi}\int_{\Gamma_1}T_{G_2}(x,f)\frac{\partial g_1(x,x_0)}{\partial n}ds=\frac{1}{2\pi}\int_{\Gamma_1}\left(T_{G_2}(x,f)-V_{G_1}(x,f)\right)\frac{\partial g_1}{\partial n}ds$$
$$=T_{G_2}(x_0,f)-V_{G_1}(x_0,f).$$

Nach (4) ist somit

$$U_{G_1}(x_0, f) \leqq T_{G_2}(x_0, f) - V_{G_1}(x_0, f),$$

oder also, gemäß der Definition (3) des Ausdrucks T_G,

$$T_{G_1}(x_0, f) \leqq T_{G_2}(x_0, f), \quad \text{w. z. b. w.}$$

3. Es sei G ein inneres Gebiet von G^*, und x_0 ein innerer Punkt von G; die Greensche Funktion des Gebietes G mit ihrem Pol in $x = x_0$ wird wie vorher mit $g(x, x_0)$ bezeichnet. Wir betrachten das von der Niveaulinie

$$g(x, x_0) = \lambda \qquad (\lambda \geqq 0)$$

begrenzte Gebiet G_λ, das für $\lambda = 0$ in G übergeht, und für $\lambda \to \infty$ sich zu dem Punkt x_0 zusammenzieht. Nach dem Satz I ist der Ausdruck $T_{G_\lambda}(x_0, f)$ eine niemals zunehmende Funktion von λ. Wir werden jetzt folgenden Zusatz beweisen, von dem wir zwar in der Folge keinen Gebrauch machen werden, der aber an sich nicht ohne Interesse sein dürfte:

Satz II. *Der Ausdruck $T_{G_\lambda}(x_0, f)$ ist eine konvexe Funktion von λ.*

Beweis. Wir wählen zwei beliebige positive Zahlen λ_1 und λ_2 $(\lambda_1 > \lambda_2)$; das von den Niveaulinien $g = \lambda_1$ und $g = \lambda_2$ begrenzte Gebiet bezeichnen wir mit H. Nach der Grundformel (1)′ ist dann in jedem Punkt x dieses Gebietes

$$\log|f(x)| = T_H(x, f) - T_H\left(x, \frac{1}{f}\right) \leqq T_H(x, f),$$

und also auch, da $T_H(x, f)$ nichtnegativ ist,

$$\overset{+}{\log}|f| \leqq T_H(x, f).$$

Diese Ungleichung besteht demnach insbesondere auf der Niveaulinie $g = \lambda$, falls $\lambda_1 \geqq \lambda \geqq \lambda_2$. Durch Multiplikation mit $\frac{\partial g(x, x_0)}{\partial n}$ und Integration längs der genannten Niveaulinie folgt also, daß

$$U_{G_\lambda}(x_0, f) \leqq \frac{1}{2\pi}\int\limits_{g=\lambda} T_H(x, f)\frac{\partial g(x, x_0)}{\partial n}\,ds. \tag{5}$$

Andererseits ergibt sich, da die Funktion $V_{G_{\lambda_2}}(x, f) - V_{G_\lambda}(x)$ innerhalb G_λ harmonisch und regulär ist und auf der Kurve $g = \lambda$ die Werte $V_{G_{\lambda_2}}(x, f)$ annimmt, daß

$$\frac{1}{2\pi}\int\limits_{g=\lambda} V_{G_{\lambda_2}}(x, f)\frac{\partial g(x, x_0)}{\partial n}\,ds = \frac{1}{2\pi}\int\limits_{g=\lambda}\left(V_{G_{\lambda_2}}(x, f) - V_{G_\lambda}(x, f)\right)\frac{\partial g}{\partial n}\,ds$$
$$= V_{G_{\lambda_2}}(x_0, f) - V_{G_\lambda}(x_0, f),$$

woraus für $V_{G_\lambda}(x_0, f)$ der Wert

$$V_{G_\lambda}(x_0, f) = V_{G_{\lambda_2}}(x_0, f) - \frac{1}{2\pi}\int\limits_{g=\lambda} V_{G_{\lambda_2}}(x, f)\frac{\partial g}{\partial n}ds$$
$$= \frac{1}{2\pi}\int\limits_{g=\lambda}(V_{G_{\lambda_2}}(x_0, f) - V_{G_{\lambda_2}}(x, f))\frac{\partial g}{\partial n}ds$$

folgt.

Wird diese Gleichung zu der Ungleichung (5) addiert, so ergibt sich

$$(6) \qquad T_{G_\lambda}(x_0, f) \leqq \frac{1}{2\pi}\int\limits_{g=\lambda} W(x)\frac{\partial g(x, x_0)}{\partial n}ds,$$

wo

$$W(x) = T_H(x, f) - V_{G_{\lambda_2}}(x, f) + V_{G_{\lambda_2}}(x_0, f).$$

Die Funktion $W(x)$ ist innerhalb H harmonisch und regulär; mittels der Greenschen Formel folgt hieraus leicht, daß der Ausdruck rechts in (6) eine *lineare* Funktion von λ ist[5]). Für $\lambda = \lambda_1$ und $\lambda = \lambda_2$ kann sein Wert leicht ermittelt werden. Auf der Kurve $g = \lambda_1$ ist nämlich

$$T_H(x, f) = \overset{+}{\log}|f| = T_{G_{\lambda_1}}(x, f),$$

und daher

$$\frac{1}{2\pi}\int\limits_{g=\lambda_1} W(x)\frac{\partial g}{\partial n}ds = \frac{1}{2\pi}\int\limits_{g=\lambda_1}(T_{G_{\lambda_1}}(x, f) - V_{G_{\lambda_2}}(x, f) + V_{G_{\lambda_2}}(x_0, f))\frac{\partial g}{\partial n}ds$$
$$= T_{G_{\lambda_1}}(x_0, f).$$

Für $g = \lambda_2$ ist wieder $T_H(x, f) = \overset{+}{\log}|f|$, $\quad V_{G_{\lambda_2}}(x, f) = 0$, und also

$$\frac{1}{2\pi}\int\limits_{g=\lambda_2} W(x)\frac{\partial g}{\partial n}ds = U_{G_{\lambda_2}}(x_0, f) + V_{G_{\lambda_2}}(x_0, f) = T_{G_{\lambda_2}}(x_0, f).$$

Die Ungleichung (6), wo die rechte Seite im Intervall $\lambda_1 \leqq \lambda \leqq \lambda_2$ eine lineare Funktion von λ ist, geht also für $\lambda = \lambda_1$ und $\lambda = \lambda_2$ in eine *Gleichheit* über. Hieraus folgt die Behauptung unmittelbar.

---

[5]) Wird nämlich die Formel $\int\left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right)ds = 0$ in dem von den Niveaulinien $g = \lambda_1$ und $g = \lambda$ begrenzten Gebiete mit $u = W$ und $v = g - \lambda$ angewandt, so folgt unmittelbar, daß

$$\int\limits_{g=\lambda} W\frac{\partial g}{\partial n}ds = c_1\lambda + c_2,$$

wo c_1 und c_2 die von λ unabhängigen Größen

$$c_1 = \int\limits_{g=\lambda_1}\frac{\partial W}{\partial n}ds, \qquad c_2 = \int\limits_{g=\lambda_1} W\frac{\partial g}{\partial n}ds - c_1\lambda_1$$

bezeichnen.

4. Wir betrachten jetzt eine unendliche Folge von Gebieten

$$G_1, G_2, \ldots, G_n, \ldots,$$

die sämtlich innere Gebiete von G^* sind, und von denen G_n für jedes n als Teilgebiet in G_{n+1} enthalten ist. Für $n \to \infty$ nähere sich die Randkurve Γ_n von G_n dem Rande Γ^* von G^* derart, daß jedes vorgegebene innere Gebiet von G^* schließlich innerhalb G_n liegt. Nach dem Satz I ist dann

$$(7) \qquad T_{G_1}(x, f) \leqq T_{G_2}(x, f) \leqq \ldots \leqq T_{G_n}(x, f) \leqq \ldots.$$

Wir wollen jetzt annehmen, es existiere innerhalb G^* ein Punkt x_0, in welchem die Folge $T_{G_n}(x, f)$ *beschränkt* ist. Dann existiert der endliche Grenzwert

$$\lim_{n=\infty} T_{G_n}(x_0, f),$$

woraus durch eine wohlbekannte Schlußweise folgt, daß die Folge (7) in *jedem* Punkt x innerhalb G^* konvergiert und zwar gleichmäßig in jedem inneren Gebiet von G^*. Der Grenzwert

$$\lim_{n=\infty} T_{G_n}(x, f) = T(x, f)$$

ist eine innerhalb G^* nichtnegative, endliche, harmonische Funktion, mit Ausnahme der Pole von $f(x)$, wo $T(x, f)$ positive, logarithmische Pole besitzt.

Wir wenden dann unseren Satz I auf die Funktion $\frac{1}{f}$ an und finden so:

$$(8) \qquad T_{G_1}\left(x, \frac{1}{f}\right) \leqq T_{G_2}\left(x, \frac{1}{f}\right) \leqq \ldots \leqq T_{G_n}\left(x, \frac{1}{f}\right) \leqq \ldots.$$

Man nehme zunächst an, daß die betrachtete Funktion $f(x)$ nicht identisch verschwindet; sei dann x_1 ein Punkt, wo $f(x) \neq 0$ und endlich ist. Die Ausdrücke (7) sind für $x = x_1$ nicht größer als die endliche Zahl $T(x_1, f)$; wendet man nun die Formel (1)′ für $x = x_1$ an, so folgt also, daß

$$T_{G_n}\left(x_1, \frac{1}{f}\right) = \log\left|\frac{1}{f(x_1)}\right| + T_{G_n}(x_1, f) \leqq \log\left|\frac{1}{f(x_1)}\right| + T(x_1, f).$$

Die Ausdrücke (8) sind somit für $x = x_1$ gleichmäßig beschränkt, woraus man schließt, daß auch der Grenzwert

$$\lim_{n=\infty} T_{G_n}\left(x, \frac{1}{f}\right) = T\left(x, \frac{1}{f}\right)$$

in jedem inneren Punkt von G^* existiert. Die Grenzfunktion $T\left(x, \frac{1}{f}\right)$ ist innerhalb G^* nichtnegativ und harmonisch, außer in den Polen von $\frac{1}{f}$ (Nullstellen von f), wo sie positive logarithmische Pole besitzt.

Mittels der Grundformel (1) ergibt sich nun, wenn man G_n statt G schreibt und n unbeschränkt wachsen läßt, für $\log|f(x)|$ der Wert

$$(9)\qquad \log|f(x)| = T(x, f) - T\left(x, \frac{1}{f}\right).$$

Es sei $\overline{T}(x, f)$ die bis auf eine reelle Konstante wohlbestimmte, zu $T(x, f)$ konjugierte harmonische Funktion, und

$$(10)\qquad \varphi(x, f) = e^{-T(x,f) - i\overline{T}(x,f)}.$$

Diese Funktion, die innerhalb G^* regulär ist, genügt der Bedingung

$$|\varphi(x, f)| \leqq 1$$

und verschwindet in den Polen von $f(x)$. Durch geeignete Wahl der in $\overline{T}$ enthaltenen willkürlichen Konstante findet man nach (9) für $f(x)$ die innerhalb G^* gültige Darstellung

$$(11)\qquad f(x) = \frac{\varphi\left(x, \frac{1}{f}\right)}{\varphi(x, f)}.$$

Dieses Resultat haben wir unter der Annahme hergeleitet, daß $f(x)$ nicht identisch verschwindet. Ist nun $f(x) \equiv 0$, so bleibt die Darstellung (11) gültig, falls man $T\left(x, \frac{1}{f}\right) \equiv +\infty$, und also $\varphi\left(x, \frac{1}{f}\right) \equiv 0$ setzt. Zusammenfassend haben wir somit folgendes Ergebnis:

Satz III. *Es sei $f(x)$ eine innerhalb eines zusammenhängenden Gebietes G^* eindeutige meromorphe Funktion, und $G_1, G_2, \ldots, G_n, \ldots$ eine Folge von inneren Gebieten, von denen G_n für jedes n als Teilgebiet in G_{n+1} enthalten ist. Die Randkurve Γ_n von G_n nähere sich mit wachsendem n unbeschränkt dem Rande von G^*, so daß schließlich jedes gegebene innere Gebiet von G^* vollständig innerhalb G_n liegt.*

Existiert innerhalb G^* ein Punkt, wo der Ausdruck $T_{G_n}(x, f)$ unter einer festen, von n unabhängigen, endlichen Schranke liegt, so konvergiert er für $n \to \infty$ in jedem inneren Punkt von G^* gegen eine harmonische Funktion $T(x, f)$. Die meromorphe Funktion $f(x)$ läßt sich dann als Quotient von zwei, durch die Formel (10) *definierten, beschränkten Funktionen $\varphi\left(x, \frac{1}{f}\right)$ und $\varphi(x, f)$ darstellen, von denen wenigstens die letztgenannte nicht identisch verschwindet.*

5. Es seien umgekehrt $\psi_1(x)$ und $\psi_2(x)$ zwei beliebige analytische Funktionen, die innerhalb G^* regulär und beschränkt sind:

$$(12)\qquad |\psi_1(x)| \leqq 1, \qquad |\psi_2(x)| \leqq 1,$$

und von denen ψ_2 nicht identisch gleich Null ist. Wir werden zeigen, daß die meromorphe Funktion

$$f(x) = \frac{\psi_1(x)}{\psi_2(x)}$$

zu der oben betrachteten Klasse gehört, d. h. daß die Folge $T_{G_n}(x, f)$ $(n = 1, 2, \ldots)$ in jedem von den Polen der Funktion $f(x)$ verschiedenen Punkt von G^* beschränkt ist.

Sei in der Tat x_0 ein Punkt, wo $\psi_2(x) \neq 0$. Nach den Bedingungen (12) ist $|f| = \left|\frac{\psi_1}{\psi_2}\right| \leqq \left|\frac{1}{\psi_2}\right|$, und also

$$\overset{+}{\log} |f(x)| \leqq \overset{+}{\log} \left|\frac{1}{\psi_2(x)}\right|.$$

Sobald die Kurve Γ_n den Punkt x_0 umfaßt, ist demnach

$$U_{G_n}(x_0, f) \leqq U_{G_n}\left(x_0, \frac{1}{\psi_2}\right).$$

Weil die Pole von $f(x)$ sämtlich in den Nullstellen von $\psi_2(x)$ enthalten sind, so wird auch

$$V_{G_n}(x_0, f) \leqq V_{G_n}\left(x_0, \frac{1}{\psi_2}\right).$$

Durch Addition der letzten zwei Ungleichungen folgt nun, daß

$$T_{G_n}(x_0, f) \leqq T_{G_n}\left(x_0, \frac{1}{\psi_2}\right). \tag{13}$$

Der Wert des letzten Ausdruckes kann leicht berechnet werden. Nach den Ungleichungen (12) ist nämlich $\overset{+}{\log} |\psi_2| \equiv 0$, und also $U_{G_n}(x_0, \psi_2) = 0$; weil $\psi_2(x)$ innerhalb G^* *regulär* ist, so verschwindet auch der Ausdruck $V_{G_n}(x_0, \psi_2)$, und es ist also

$$T_{G_n}(x_0, \psi_2) = 0.$$

Mittels der Formel $(1)'$ ergibt sich nun für die rechte Seite von (13) der endliche Wert

$$T_{G_n}\left(x_0, \frac{1}{\psi_2}\right) = \log \left|\frac{1}{\psi_2(x_0)}\right|.$$

Man sieht also, daß die Ungleichung

$$T_{G_n}(x_0, f) \leqq \log \left|\frac{1}{\psi_2(x_0)}\right| \tag{14}$$

für jedes n besteht, womit die Behauptung nachgewiesen ist. Wir haben also den

Satz IV. *Notwendig und hinreichend, damit eine innerhalb G^* meromorphe Funktion sich als Quotient von zwei beschränkten Funktionen darstellen läßt, ist, daß ein innerer Punkt existiert, wo die Folge $T_{G_n}(x, f)$ $(n = 1, 2, \ldots)$ gleichmäßig beschränkt ist.*

6. Wenn eine meromorphe Funktion zu der oben betrachteten Klasse gehört, so kann sie selbstverständlich in unendlich vielen Weisen als Quotient von zwei beschränkten Funktionen dargestellt werden. Unter

diesen Darstellungen zeichnet sich die in (11) gegebene, wo $\varphi(x, f)$ durch die Formel (10) definiert ist, durch folgende Eigenschaft aus:

Satz V. *Es seien $\psi_1(x)$ und $\psi_2(x)$ zwei innerhalb G^* reguläre, beschränkte Funktionen* ($|\psi_1| \leqq 1$, $|\psi_2| \leqq 1$), *und*

$$f(x) = \frac{\psi_1(x)}{\psi_2(x)}.$$

Dann gilt in jedem inneren Punkt von G

$$|\varphi(x, f)| \geqq |\psi_2(x)|, \qquad \left|\varphi\left(x, \frac{1}{f}\right)\right| \geqq |\psi_1(x)|.$$

Der Nachweis ergibt sich unmittelbar aus der oben bewiesenen Ungleichung (14), die für jeden inneren Punkt x_0 des Gebietes G besteht. Weil nämlich $\lim\limits_{n=\infty} T_{G_n}(x, f) = T(x, f)$, so ist demnach auch

$$T(x, f) \leqq \log\left|\frac{1}{\psi_2(x)}\right|,$$

oder also, nach der Definition (10),

$$|\varphi(x, f)| \geqq |\psi_2(x)|.$$

Der die Funktion $\varphi\left(x, \frac{1}{f}\right)$ betreffende Teil folgt, wenn man dieselbe Schlußweise auf die Funktion $\frac{1}{f(x)}$ anwendet.

Aus dem Satz V geht insbesondere hervor, daß die zu einer meromorphen Funktion gehörigen beschränkten Funktionen $\varphi(x, f)$ und $\varphi\left(x, \frac{1}{f}\right)$ *eindeutig* bestimmt sind, d. h. daß sie von der besonderen Wahl der Gebietfolge G_n unabhängig sind, wenn diese nur den im Satz III angegebenen Bedingungen genügt.

7. Um zu dem in der Einleitung angegebenen Satz zu gelangen, wollen wir zum Schluß die oben erzielten Ergebnisse auf den Fall anwenden, wo das Gebiet G^* ein Kreis $|x| < R$ ist. Wählen wir als Gebiet G_n einen konzentrischen Kreis C_ϱ vom Radius ϱ ($\varrho < R$), so wird für $x = re^{i\varphi}$ ($r < \varrho$)

$$(15) \qquad T_{C_\varrho}(x, f) = \frac{1}{2\pi}\int_0^{2\pi} \overset{+}{\log}|f(\varrho e^{i\vartheta})| \frac{\varrho^2 - r^2}{\varrho^2 + r^2 - 2\varrho r\cos(\vartheta - \varphi)}\, d\vartheta$$
$$+ \sum_{|b_\nu| < \varrho} \log\left|\frac{\varrho^2 - \bar{b}_\nu x}{\varrho(x - b_\nu)}\right|,$$

wo b_ν die Pole von $f(x)$ bezeichnen. Der in der Einleitung angegebene Satz ergibt sich nun aus den oben bewiesenen Sätzen I, II, III und IV,

wenn man in (15) insbesondere $x = 0$ wählt und bemerkt, daß dieser Ausdruck durch eine leichte Umformung auf die Form

$$(16) \qquad \frac{1}{2\pi}\int_0^{2\pi} \overset{+}{\log} |f(\varrho e^{i\vartheta})|\, d\vartheta + \int_0^{\varrho} \frac{n(r)}{r}\, dr$$

gebracht werden kann, wobei $n(r)$ die Anzahl der Pole von $f(x)$ für $|x| < r$ bezeichnet.

Man sieht leicht ein, daß das Integral

$$\int_0^R \frac{n(r)}{r}\, dr$$

mit der Reihe

$$(17) \qquad \sum_{\nu=1}^{\infty} (R - |b_\nu|)$$

gleichzeitig konvergent oder divergent ist. Das genannte Integral ist nämlich in bezug auf Konvergenz und Divergenz mit dem Integral

$$\int_0^R n(r)\, dr$$

gleichwertig; und aus der Identität

$$\int_0^{\varrho} n(r)\, dr = \int_0^{\varrho} n(r)\, d(r - R) = -(R - \varrho)\, n(\varrho) + \sum_{|b_\nu| < \varrho} (R - |b_\nu|)$$

ersieht man unmittelbar, daß das linksstehende Integral mit der Reihe (17) gleichzeitig konvergent oder divergent ist.

Das in der Einleitung angegebene Kriterium (*Beschränktheit des Ausdruckes* (16)) *ist also damit gleichbedeutend, daß das Integral*

$$\frac{1}{2\pi}\int_0^{2\pi} \overset{+}{\log} |f(re^{i\varphi})|\, d\varphi$$

für $r < R$ *beschränkt und die Reihe* (17) *konvergent ist.*

(Eingegangen am 3. 2. 1924.)

Über indefinite binäre quadratische Minimalformen.

Von

Robert Remak in Berlin.

Einleitung.

Frobenius faßt in seiner Arbeit „Über die Markoffschen Zahlen"[1]) (zitiert F. M.) die Resultate der beiden Arbeiten des Herrn Markoff „Sur les formes quadratiques binaires indéfinies"[2]) folgendermaßen zusammen:

„In der indefiniten Form

$$\psi = (a, b, c) = ax^2 + bxy + cy^2$$

seien die Koeffizienten a, b, c beliebige reelle Größen, die Variablen x, y ganze Zahlen. Die Diskriminante von ψ sei $D = b^2 - 4ac$, die untere Schranke aller Werte des absoluten Betrages $|\psi|$ von ψ sei M (außer für $x = 0$, $y = 0$). Für die Formen $k\psi$ haben diese Größen die Werte $k^2 D$ und kM. Dann beweist Herr Markoff:

Für die Gesamtheit aller indefiniten Formen ψ ist

$$\liminf \frac{\sqrt{D}}{M} = 3.$$

Ist $\sqrt{D} < 3M$, so ist ψ, mit einem passenden Faktor k multipliziert, einer Form

$$\varphi = px^2 + (3p - 2q)xy + (r - 3q)y^2 \tag{1}$$

(*eigentlich oder uneigentlich*) *äquivalent. Hier sind p, q, r positive ganze Zahlen; p genügt, zusammen mit zwei anderen Zahlen, p_1 und p_2, der unbestimmten Gleichung*

$$p^2 + p_1^2 + p_2^2 = 3pp_1p_2, \tag{2}$$

[1]) Sitzungsber. d. Preuß. Akad. d. Wissensch. (1913), S. 458–487.

[2]) Math. Annalen 15, S. 381–409 und 17, S. 379–399.

$\pm q$ ist der absolut kleinste Rest von $\frac{p_1}{p_2}$ $(\mathrm{mod}\, p)$ *und r ist durch die Gleichung $pr - q^2 = 1$ bestimmt. Für diese Form φ ist*

$$D = 9p^2 - 4, \quad M = p, \quad \frac{\sqrt{D}}{M} = 3\sqrt{1 - \frac{4}{9p^2}} < 3.$$

Die Form φ ist der Form $-\varphi$ eigentlich äquivalent, und sogar jeder der vier Formen $\pm(px^2 - 2qxy + ry^2) \pm 3y(px - qy)$. Sind die Verhältnisse der Koeffizienten von ψ nicht rational, so ist stets $M \leqq \frac{1}{3}\sqrt{D}$."

Während Herr Markoff seine Beweise durch sehr mühsame Kettenbruchuntersuchungen führt, leitet Frobenius ohne Benutzung von Kettenbrüchen die angegebenen Eigenschaften der Form φ ab. Es gelingt ihm aber nicht, ohne dies Hilfsmittel zu beweisen, daß die mit den φ äquivalenten Formen die einzigen sind, für die $\sqrt{D} < 3M$ ist. Diese Lücke will ich in dieser Arbeit ausfüllen. Von den Formeln, die Frobenius im § 2 seiner Arbeit (F. M.) aufstellt, werde ich ausgiebigen Gebrauch machen, auch die in § 8 angegebene Zuordnung der Markoffschen Zahlen und der rationalen Brüche benutzen, dagegen keinen Gebrauch machen von den Kettenbruchentwicklungen für die Markoffschen Zahlen.

In Anlehnung an Herrn I. Schur[3]) will ich eine binäre quadratische Form $f(x, y) = ax^2 + bxy + cy^2$ dann eine (gewöhnliche) Minimalform nennen, wenn für alle ganzzahligen x, y außer 0, 0

$$(3)\,(\mathrm{A}) \qquad |f(x, y)| \geqq |a|.$$

Es gibt also stets unendlich viele parallele Minimalformen. Ist die Form indefinit, so gibt es unter diesen stets eine reduzierte Form; diese heiße reduzierte Minimalform. Die Bedingung dafür, daß eine indefinite Form (a, b, c) reduziert sei, lautet[4])

$$(4) \qquad b < \sqrt{D}, \quad b > \sqrt{D} - 2|a|, \quad b > \sqrt{D} - 2|c|.$$

Hierin bedeutet $D = b^2 - 4ac$ die Diskriminante der Form a, b, c. Da man, ohne a zu ändern, durch eine Paralleltransformation b um beliebige

---

[3]) „Zur Theorie der indefiniten binären quadratischen Formen". Sitzungsber. d. Preuß. Akademie d. Wissensch. (1913), S. 212–231. Herr Schur definiert als Minimalform $\mu = (a, b, -c)$, in der $0 \leqq b \leqq a \leqq c$, $a > 0$, und a zugleich die kleinste durch $|\mu|$ darstellbare Zahl. Derartige Formen, die einen speziellen Fall der gewöhnlichen Minimalformen bilden, mögen normale Minimalformen heißen. Für negative Diskriminanten stimmen die normalen Minimalformen mit den reduzierten Formen überein. Jede definite quadratische Form erreicht ihre untere Grenze des absoluten Betrages, ist also einer Minimalform äquivalent.

[4]) Siehe Frobenius, „Über die Reduktion der indefiniten binären quadratischen Formen" (zitiert mit F. R.). Sitzungsber. d. Preuß. Akad. d. Wissensch. (1913), § 1, S. 203 (5).

ganzzahlige Vielfache von $2a$ verändern kann, so lassen sich die erste und zweite Bedingung (4) durch ein bestimmtes b erfüllen, die dritte ist dann von selbst erfüllt, weil nach (3)(A)

(5)(a) $$|c| \geqq |a|.$$

Da nicht jede indefinite quadratische Form die untere Grenze ihres absoluten Betrages erreicht, so ist nicht jede indefinite quadratische Form einer Minimalform äquivalent.

Von den drei Behauptungen:

I. *Die Formen* (1) *sind Minimalformen.*

II. *Die Formen* (1) *und die ihnen proportionalen Formen sind die einzigen Minimalformen, für die*

(6) $$\frac{\sqrt{D}}{|a|} < 3.$$

III. *Eine Form, die keiner Form* (1) *oder einer proportionalen äquivalent ist, ist äquivalent einer Form, für die*

(7) $$\frac{\sqrt{D}}{|a|} > 3 - \varepsilon,$$

wo ε eine beliebig kleine positive vorgegebene Zahl,

hat Frobenius I. in einfacher Weise bewiesen (F. M. § 4, S. 467), II. ist ein spezieller Fall von III. Der Übersichtlichkeit halber werde ich in § 1 dieser Arbeit vorweg II. beweisen, obwohl sich die Trennung auch vermeiden ließe; dabei wird schon ein Teil der Beziehungen aufgestellt, die zum Beweise von III. nötig sind; dieser wird in § 2 gegeben. § 3, der sich nur auf § 1 stützt, wird endlich eine vollständige Übersicht über die Minimalform enthalten, für die

(8) $$\frac{\sqrt{D}}{a} = 3.$$

Da es nur auf den Quotienten $\frac{\sqrt{D}}{a}$ ankommt, so darf man jede Form mit einer beliebigen reellen Zahl multipliziert denken. Insbesondere bleibt bei Multiplikation mit -1 die Diskriminante D ungeändert. Es werde also im folgenden stets, wenn nicht das Gegenteil ausdrücklich bemerkt ist, $a > 0$ angenommen. Der Übersichtlichkeit halber werde ich bei Ungleichheitsbeziehungen von der Form (3) der laufenden Nummer der Beziehung ein (A) (absoluter Betrag) hinzufügen.

(3)(A) $$|f(x, y)| \geqq a > 0$$

zerfällt in die beiden Unterfälle

(3)(P) $$f(x, y) \geqq a$$

und

(3)(N) $$f(x, y) \leqq -a$$

((P) positiv, (N) negativ). Wird anstatt $f(x, y)$ explizit $ax^2 + bxy + cy^2$ geschrieben, so treten an die Stelle der großen Buchstaben die kleinen Buchstaben (a), (p), (n).

§ 1.

Für eine reduzierte Minimalform

(9) $$f(x, y) = ax^2 + bxy + cy^2 = (a, b, c)$$

sei

(10) $$a > 0,$$

dann ist nach F. R. § 1, S. 204 (13)

$$c < 0$$

(11) $$b > |a + c| = |c| - a$$

mit Rücksicht auf (5)(a). Die Bedingung

(12)(A) $$|f(1, 1)| \geqq a$$

zerfällt in die Unterfälle

(12)(p) $a + b + c \geqq a$ oder (13)(p) $b \geqq |c|$

und

(12)(n) $a + b + c \leqq -a$ oder (13)(n) $b \leqq |c| - 2a.$

(13)(n) widerspricht (11); für reduzierte Minimalformen gilt also stets die Bedingung (13)(p), die schärfer ist als (11). Für reduzierte Minimalformen gilt also die (13)(p) und (5)(a) zusammenfassende Bedingung

(14) $$b \geqq |c| \geqq a.$$

Die Bedingung

(15)(A) $$|f(-1, 1)| \geqq a$$

liefert

(15)(n) $$a - b + c \leqq -a,$$

die durch (14) von selbst erfüllt ist. Nimmt man die Bedingung

(16)(A) $$|f(-2, 1)| \geqq a$$

hinzu, so kann

(16)(p) $$4a - 2b + c \geqq a$$

sein oder in Verbindung mit (14)

(17)(p) $\quad 3a \geqq 2b + |c| \geqq 3|c| \quad$ folglich $\quad a \geqq |c|$,

also wegen (14) $a = |c|$.

Setzt man das in (17)(p) ein, so erhält man $a \geqq b$ und wegen (14) $a = b$. Also handelt es sich um die spezielle Form

$$ax^2 + axy - ay^2, \tag{18}$$

die wirklich eine Minimalform ist, weil sie nur ganzzahlige Vielfache von a darzustellen vermag. Für alle von (18) verschiedenen reduzierten Minimalformen gilt also

(16)(n) $$4a - 2b + c \leqq -a.$$

Man nehme die beiden Bedingungen

(19)(A) $\quad |f(1,2)| \geqq a, \qquad$ (20)(A) $\quad |f(-5,2)| \geqq a$

hinzu. Wenn gleichzeitig

(19)(p) $\quad a + 2b + 4c \geqq a \quad$ und $\quad$ (20)(p) $\quad 25a - 10b + 4c \geqq a$,

so multipliziere man (19)(p) beiderseits mit 5, addiere dazu (20)(p) und erhält $a \geqq |c|$, also in Verbindung mit (14) $a = |c|$. Das gibt, in (19)(p) eingesetzt, $b \geqq 2a$, in (20)(p) eingesetzt, $2a \geqq b$, also ist $b = 2a$. Es liegt also die spezielle reduzierte Form

$$ax^2 + 2axy - ay^2 \tag{21}$$

vor, die wirklich eine Minimalform ist, weil sie nur ganzzahlige Vielfache von a darzustellen vermag. Für alle von (18) und (21) verschiedenen reduzierten Formen ist also mindestens eine der beiden Bedingungen

(19)(N) $\quad f(1,2) \leqq -a, \qquad$ (20)(N) $\quad f(-5,2) \leqq -a$

erfüllt. Man stelle

(19)(N) $\quad f(1,2) \leqq -a \quad$ und $\quad$ (16)(N) $\quad f(-2,1) \leqq -a$

einerseits,

(5)(N) $\quad f(0,1) \leqq -a \quad$ und $\quad$ (20)(N) $\quad f(-5,2) \leqq -a$

andererseits zusammen. Jede von (18) und (21) verschiedene reduzierte Minimalform genügt dem einen oder dem anderen Bedingungspaare. Durch die Substitution

$$\begin{cases} x = -x' - 2y', \\ y = y', \end{cases} \tag{22}$$

die zu sich selbst reziprok ist, mit der Determinante -1 geht (19)(N) in (20)(N) und (16)(N) in (5)(N), jede Minimalform wieder in eine

uneigentlich äquivalente Minimalform über. Ich kann mich also im folgenden auf die Untersuchung der Minimalformen beschränken, die (20) (N) und (5) (N) genügen.

Der Beweis wird von nun an durch Induktion geführt. Es seien p, p_1, p_2 drei konjugierte Markoffsche Zahlen, die (2) erfüllen. p sei die größte unter ihnen; q, q_1, q_2 seien durch die Kongruenzen F. M., § 2, S. 462, (2) definiert und es sei *ständig* für das folgende

(23) $$\varepsilon = -1$$

gesetzt: Es möge eine Menge von Minimalformen den beiden Bedingungen genügen

(24) (N) $f(q_1, p_1) \leqq -a,$ (25) (N) $f(q_2 - 3p_2, p_2) \leqq -a.$

Multipliziert man (24) (N) mit der positiven Zahl $p_2(3p_2 - q_2)$, (25) (N) mit $p_1 q_1$, addiert man die Ungleichheitsbeziehungen, so fällt b heraus, und man erhält unter Benutzung der Formeln F. M., § 2, (4), (5), (14) (17) und (4)

(26) $$a(3q - r) + pc \leqq 0.$$

Man nehme die beiden Bedingungen

(27) (A) $|f(q, p)| \geqq a,$ (28) (A) $|f(q - 3p, p)| \geqq a$

hinzu. Man betrachte zunächst den Fall, daß sowohl

(27) (P) $f(q, p) \geqq a$ als auch (28) (P) $f(q - 3p, p) \geqq a$

besteht. Multipliziert man (27) (P) mit der positiven Zahl $3p - q$, (28) (P) mit q, addiert man die Ungleichheitsbeziehungen, so fällt b heraus und man erhält unter Benutzung von F. M., § 2, (3)

$$a(3q - r) + cp \geqq 0.$$

Hieraus folgt in Verbindung mit (26)

(29) $$c = a\frac{r - 3q}{p}.$$

Setzt man diesen Wert in (24) (N) ein, so erhält man unter Benutzung der Formeln F. M., § 2, (3), (14) und (4)

(30) $$a(2q - 3p) + bp \leqq 0.$$

Setzt man ferner (29) in (25) (N) ein, so erhält man unter Benutzung derselben Formeln und unter Fortlassung des negativen Faktors $q_2 - 3p_2$:

$$a(2q - 3p) + bp \geqq 0.$$

Hieraus folgt in Verbindung mit

(30) $$b = a\frac{3p - 2q}{p}.$$

Also handelt es sich um die spezielle Form

$$(31)\qquad \frac{a}{p}(px^2+(3p-2q)x+(r-3q)y^2)=\frac{a}{p}\varphi(x,y),$$

eine zu (1) proportionale Form, die nach der von Frobenius bewiesenen Behauptung I. eine Minimalform ist. Alle von (31) verschiedenen Minimalformen, die (24)(N) und (25)(N) genügen, genügen also mindestens einer der beiden Bedingungen

$$(27)(\mathrm{N})\quad f(q,p)\leqq -a,\qquad\qquad (28)(\mathrm{N})\quad f(q-3p,p)\leqq -a.$$

Man stelle

$$(27)(\mathrm{N})\quad f(q,p)\leqq -a\quad \text{und}\quad (25)(\mathrm{N})\quad f(q_2-3p_2,p_2)\leqq -a$$

einerseits,

$$(24)(\mathrm{N})\quad f(q_1,p_1)\leqq -a\quad \text{und}\quad (28)(\mathrm{N})\quad f(q-3p,p)\leqq -a$$

andererseits zusammen. Es ist somit bewiesen worden: Alle Minimalformen, die einem Paare von Bedingungen von der Form (24)(N), (25)(N) genügen, genügen mit Ausnahme der einzigen Minimalform (31) mindestens einem der beiden Bedingungspaare (27)(N), (25)(N) oder (24)(N), (28)(N), die von derselben Form wie (24)(N), (25)(N) sind, aber größere Markoffsche Zahlen enthalten.

Läßt man speziell in (24)(N), (25)(N) das Gleichheitszeichen gelten, so ist dadurch die Form f bis auf einen Proportionalitätsfaktor vollständig bestimmt, und zwar gleich der Form (31). Es gelten die Identitäten, die sich auch leicht direkt verifizieren lassen:

$$(32)\quad \varphi(q_1,p_1)=-p,\qquad\qquad (33)\quad \varphi(q_2-3p_2,p_2)=-p,$$

$$(34)\quad \varphi(q,p)=p,\qquad\qquad (35)\quad \varphi(q-3p,p)=p.$$

Es soll nun die untere Grenze von $\frac{D}{a^2}$ bestimmt werden für diejenigen Formen, für die (24)(N) und (25)(N) gelten. Unter Benutzung von F. M., § 2 (3) lassen sich (24)(N) und (25)(N) in der Form schreiben

$$(36)(\mathrm{n})\qquad a\frac{r_1}{p_1}+b\frac{q_1}{p_1}+c\leqq 0$$

$$(37)(\mathrm{n})\qquad a\left(\frac{r_2}{p_2}-6\frac{q_2}{p_2}+9\right)+b\left(\frac{q_2}{p_2}-3\right)+c\leqq 0.$$

Da alle $\frac{q}{p}$ echte Brüche sind, so ist

$$\frac{q_1}{p_1}>\frac{q}{p}-\frac{3}{2}>\frac{q_2}{p_2}-3.$$

Man kann also zwei *positive* Zahlen λ_1 und λ_2 so bestimmen, daß

$$\lambda_1+\lambda_2=1,$$

$$\lambda_1\frac{q_1}{p_1}+\lambda_2\left(\frac{q_2}{p_2}-3\right)=\frac{q}{p}-\frac{3}{2}.$$

Man multipliziere (36) (n) mit λ_1, (37) (n) mit λ_2, addiere die Ungleichheitsbeziehungen und bezeichne den Faktor von a vorläufig mit x

$$ax + b\left(\frac{q}{p} - \frac{3}{2}\right) + c \leqq 0. \tag{38}$$

Für die spezielle Form φ gilt in (24) (N) und (25) (N) oder in (36) (n) und (37) (n), also auch in (38) das Gleichheitszeichen. Also wird

$$px + (3p - 2q)\left(\frac{q}{p} - \frac{3}{2}\right) + (r - 3q) = 0$$

oder

$$4p^2x = 2(3p - 2q)^2 + 12pq - 4pr = (3p - 2q)^2 + 9p^2 - 4. \tag{39}$$

Multipliziert man (38) mit $4ap^2$ und setzt man (39) ein, so erhält man

$$\{a(3p - 2q) - bp\}^2 + a^2(9p^2 - 4) - Dp^2 \leqq 0,$$

also wird

$$\frac{D}{a^2} \geqq 9 - \frac{4}{p^2}. \tag{40}$$

Für die Form (31) gilt in (40) das Gleichheitszeichen. Für alle von (31) verschiedenen Minimalformen, die (24) (N) und (25) (N) genügen, treten, wenn sie (27) (N) und (25) (N) genügen, an die Stelle von p, p_1, p_2 die drei Markoffschen Zahlen p_1', p, p_2 deren größte p_1'. Für sie ist also

$$\frac{D}{a^2} \geqq 9 - \frac{4}{p_1'^2}. \tag{41}$$

Wenn aber (24) (N) und (28) (N) gelten, so treten an die Stelle von p, p_1, p_2 die drei Markoffschen Zahlen p_2', p_1, p, deren größte p_2'. Für diese Formen wird also

$$\frac{D}{a^2} \geqq 9 - \frac{4}{p_2'^2}. \tag{42}$$

Es ist aber

$$p_1' = 3pp_2 - p_1 \geqq 2p, \quad p_2' = 3pp_1 - p_2 \geqq 2p,$$

also folgt sowohl aus (41) wie aus (42)

$$\frac{D}{a^2} \geqq 9 - \frac{4}{(2p)^2}.$$

Durch wiederholte Anwendung dieser Schlüsse ergibt sich, daß abgesehen von endlich vielen Ausnahmen für alle Minimalformen, die (24) (N) und (25) (N) genügen,

$$\frac{D}{a^2} \geqq 9 - \frac{4}{(2^n p)^2},$$

wo n eine beliebige positive ganze Zahl. Setzt man speziell

$$p_1 = 1, \quad q_1 = 0, \quad r_1 = 1, \tag{43}$$

$$p_2 = 2, \quad q_2 = 1, \quad r_2 = 1, \tag{44}$$

so ergibt sich aus den Formeln F. M., § 2, (4)

(45) $$p = 5, \quad q = 2, \quad r = 1,$$

und so geht (24)(N) in (5)(N), (25)(N) in (20)(N) über. Es genügen also alle Minimalformen mit endlich vielen Ausnahmen der Bedingung

$$\frac{D}{a^2} \geqq 9 - \frac{4}{(2^n \cdot 5)^2}$$

oder der Bedingung

(46) $$\frac{D}{a^2} \geqq 9 - \varepsilon,$$

wo ε eine beliebig kleine positive Zahl. Damit ist gezeigt, daß die zu (1) proportionalen Formen die einzigen Minimalformen sind, für die (6) erfüllt ist, also Satz II bewiesen.

Frobenius zeigt (F. M., § 5, S. 468), daß in den Formeln F. M., § 2, (4), wenn man ε festhält und z. B. p_1 und p_2 vertauscht, q in $p - q$ übergeht. Derselbe Übergang findet statt, wenn man z. B. ε, p, p_2 festhält und p_1 durch p_1' ersetzt. Ersetzt man also wie oben unter Festhaltung von (23) p, p_1, p_2 durch p_1', p, p_2 oder durch p_2', p_1, p, so bleibt q ungeändert. Ferner zeigt Frobenius, daß die drei Brüche

$$\frac{q}{p}, \quad \frac{q_1}{p_1}, \quad \frac{q_2}{p_2}$$

entweder alle kleiner als $\frac{1}{2}$ oder größer als $\frac{1}{2}$ sind. Da sie für die speziellen Werte (43), (44), (45) kleiner als $\frac{1}{2}$ sind, so sind sie kleiner als $\frac{1}{2}$ für alle daraus abgeleiteten Formen (1), d. h. es ist

(47) $$2q < p$$

für alle Minimalformen (1), die den Bedingungen (5)(N) und (20)(N) genügen.

Die Substitution (22) hat zur Folge

(48) $$a' = a, \quad b' = 4a - b, \quad c' = 4a - 2b + c.$$

Setzt man speziell

$$a = p, \quad b = 3p - 2q, \quad c = r - 3q,$$
$$a' = p', \quad b' = 3p' - 2q', \quad c' = r' - 3q',$$

in (48) ein, so erhält man

$$p' = p, \quad q' = p - q, \quad r' = p - 2q + r.$$

Für die Minimalformen von der Form (1), die (19)(N) und (16)(N) genügen, ist also $\frac{q}{p} > \frac{1}{2}$.

§ 2.

Es soll nunmehr die Behauptung III bewiesen werden. Gegeben sei eine indefinite binäre quadratische Form (9), die (10) genüge, für kein von 0, 0 verschiedenes Paar ganzer Zahlen x, y verschwinde, und weder einer Form (1) noch einer proportionalen Form äquivalent sei. Es soll eine Substitution so angegeben werden, daß die transformierte Form der Bedingung (7) oder der gleichbedeutenden Bedingung (46) genügt. Diese Substitution wird zusammengesetzt aus einer Reihe passend gewählter Substitutionen, deren Schema dem fortgesetzt in je zwei Unterfälle gegabelten der Markoffschen Minimalformen entspricht. Die ersten Schritte dieses Verfahrens, bei denen Symmetrien benutzt werden, die später fortfallen, mögen zahlenmäßig durchgeführt werden.

Sollte zunächst $b^2 > D$ sein, also $4ac = b^2 - D > 0$, a und c beide positiv, so wende man die Substitution $x = kx' + y'$, $y = -x'$ mit der Determinante 1 an, aus der

$$c' = a, \qquad b' = 2ka - b$$

folgt. Man bestimme k so, daß

$$|b'| \leqq a. \tag{49}$$

Sollte $b'^2 > D$ sein, so ist $0 < 4a'a = b'^2 - D < b'^2 \leqq a^2$, also $0 < 4a' < a$. Wendet man dieses Verfahren wiederholt an, so wird

$$0 < 4^n a^{(n)} < 4^{n-1} a^{(n-1)} < \ldots < 4a' < a.$$

Also wird nach einer endlichen Anzahl von Schritten erreicht, daß $a^{(n)2} \leqq D$, also wegen (49)

$$b^{(n+1)2} \leqq a^{(n)2} \leqq D, \qquad 4a^{(n+1)} c^{(n+1)} = b^{(n+1)2} - D \leqq 0,$$

mithin haben $a^{(n+1)}$ und $c^{(n+1)}$ entgegengesetztes Vorzeichen. Man schreibe die transformierte Form wieder (a, b, c). Es ist also (nötigenfalls durch Multiplikation mit -1) erreicht, daß

$$a > 0, \quad c < 0. \tag{50}$$

Sollte $|c| \leqq a$ sein, so wende man die Substitution $x = y'$, $y = -x'$ an und erhält $a' = c$, $b' = -b$, $c' = a$, also $|c'| \geqq |a'|$. Durch Multiplikation mit -1 erreicht man, daß (50) wieder erfüllt ist. Es ist also stets erreichbar, daß (10) erfüllt ist und

$$|c| \geqq a \quad \text{oder} \quad c \leqq -a. \tag{51}$$

Sollte $b < 0$ sein, so wende man die Substitution $x' = x$, $y' = -y$

mit der Determinante -1 an, die b in $-b$ überführt. Es ist also außer (10) und (51) noch erreichbar, daß

$$b \geqq 0. \tag{52}$$

Wenn aber (10), (51), (52) erfüllt sind, so ist

$$f(-1, 1) = a - b + c \leqq 0.$$

Sollte noch nicht

$$f(-1, 1) \leqq -a \tag{53}$$

sein, d. h. sollte

$$-a < f(-1, 1) < 0 \tag{54}$$

sein, so wende man die Substitution

$$\begin{aligned} x &= -x' - 2y', \\ y &= x' + y' \end{aligned}$$

mit der Determinante 1 an. Es wird nach (54)

$$-a < f(-1, 1) = a' < 0. \tag{55}$$

Ferner wird $f'(-1, 1) = f(1, 0) = a$, also

$$\frac{f'(-1, 1)}{a'} = \frac{a}{f(-1, 1)} < -1.$$

Es ist also $a' < 0$; es soll gezeigt werden, daß $c' > 0$ und daß $c' > |a'|$. Es genügt also zu zeigen, daß

$$c' + a' > 0. \tag{56}$$

Nun ist aber nach (55) und (51)

$$c' + a' = f(-2, 1) + f(-1, 1) = 3(f(-1, 1) + a) - (c + a) > 0,$$

also ist auch (56) erfüllt. Es ist mithin (nach Multiplikation mit -1) stets erreichbar, daß (10), (53) und (51) gleichzeitig erfüllt sind. Sollte

$$f(1, 1) \geqq f(-2, 1) \tag{57}$$

nicht erfüllt sein, so wende man die Substitution

$$\left\{ \begin{aligned} x &= -x' - y', \\ y &= \phantom{-x' -{}} y' \end{aligned} \right. \tag{58}$$

mit der Determinante -1 an, die zu sich selbst invers ist, dann wird

$$a' = a, \qquad c' = f(-1, 1), \qquad f'(-1, 1) = f(0, 1) = c.$$

Also geht durch (58) jede der Beziehungen (51) und (53) in die gestrichene andere über. Ferner folgt aus (58)

$$f'(1, 1) = f(-2, 1), \qquad f'(-2, 1) = f(1, 1);$$

ist also (57) nicht erfüllt, so ist die gestrichene Beziehung (57) erfüllt. Ich nehme also im folgenden an, es seien (10), (51), (53) und (57) erfüllt.

Durch die Identität

$$(59) \qquad 2f(1,1)+f(-2,1)-6a=3c$$

geht (51) über in

$$(60) \qquad 2f(1,1)+f(-2,1)\leqq 3a;$$

durch die Identität

$$(61) \qquad f(1,1)+2f(-2,1)-6a=3f(-1,1)$$

geht (53) über in

$$(62) \qquad f(1,1)+2f(-2,1)\leqq 3a.$$

Da aber (57) besteht, so ist (62) eine Folge von (60) oder (53) eine Folge von (51). Ferner folgt aus (57) und (60)

$$3f(-2,1)\leqq 2f(1,1)+f(-2,1)\leqq 3a,$$

also

$$(63) \qquad f(-2,1)\leqq a.$$

Die Substitution

$$(64) \qquad \begin{cases} x=2x'-y', \\ y=-x'+y' \end{cases}$$

mit der Determinante 1 ergibt

$$(65) \quad a'=f(2,-1)=f(-2,1), \qquad (66) \quad c'=f(-1,1)$$

$$(67) \quad f'(1,1)=f(1,0)=a, \qquad (68) \quad f'(-2,1)=f(-5,3).$$

Es gilt die Identität

$$(69) \qquad f(-5,3)=8(f(-2,1)-a)+f(1,1).$$

Ferner folgt aus (60) und (63)

$$(70) \quad f(1,1)-a\leqq -\tfrac{1}{2}(f(-2,1)-a)\leqq -(f(-2,1)-a).$$

Es ist nach (68), (67), (69), (70), (63)

$$f'(-2,1)-f'(1,1)=f(-5,3)-a$$
$$=8(f(-2,1)-a)+(f(1,1)-a)\leqq 7(f(-2,1)-a)\leqq 0,$$

es gilt also die gestrichene Beziehung (57). Es ist ferner nach (66), (53), (63), (65)

$$c'=f(-1,1)\leqq -a\leqq -f(-2,1)=-a',$$

also gilt die gestrichene Beziehung (51). Aus dieser folgt aber in Verbindung mit der gestrichenen Beziehung (57) das Bestehen der gestrichenen Beziehung (53), wie mit Hilfe von (59) bis (62) gezeigt wurde. Solange nun

$$a' = f(-2, 1) > 0,$$

folgt, da wegen (63)

$$a' \leqq a,$$

aus (68), (69), (70), (65), (63)

$$\frac{f'(-2,1)-a'}{a'} = \frac{f(-5,3)-f(-2,1)}{a'}$$

$$= \left\{7\frac{f(-2,1)-a}{a} + \frac{f(1,1)-a}{a}\right\}\frac{a}{a'} \leqq 6\frac{f(-2,1)-a}{a} \leqq 0.$$

Wenn also nicht zufällig

$$f(-2, 1) = a, \tag{71}$$

so tritt an die Stelle des negativen Ausdrucks

$$\frac{f(-2,1)-a}{a}$$

durch die Substitution (64) ein negativer, absolut mindestens 6 mal so großer Ausdruck. Durch wiederholte Anwendung von (64) läßt sich also nach einer endlichen Anzahl von Schritten erreichen, daß

$$\frac{f^{(n)}(-2,1)-a^{(n)}}{a^{(n)}} \leqq -1 \quad \text{oder} \quad \frac{f^{(n)}(-2,1)}{a^{(n)}} \leqq 0.$$

Es ist also stets erreichbar, daß

$$\frac{f(-2,1)}{a} \leqq 0.$$

Sollte damit noch nicht erreicht sein, daß

$$f(-2, 1) \leqq -a, \tag{72}$$

wenn also die Bedingungen (10), (51) und

$$-a < f(-2, 1) < 0 \tag{73}$$

gelten, so wende man die Substitution

$$\left\{\begin{aligned} x &= -2x' - 5y', \\ y &= x' + 2y' \end{aligned}\right. \tag{74}$$

mit der Determinante 1 an und erhält nach (73)

$$\frac{f'(-2,1)}{a'} = \frac{f(-1,0)}{f(-2,1)} = \frac{a}{f(-2,1)} < -1.$$

Ferner folgt aus (74), (73), (63), (51)

$$\frac{c'}{a'}=\frac{f(-5,2)}{f(-2,1)}=\frac{5(f(-2,1)+a)-c}{f(-2,1)}<\frac{-c}{f(-2,1)}\leqq\frac{a}{f(-2,1)}<-1.$$

Nach Multiplikation mit -1 bestehen also die gestrichenen Beziehungen (10), (72), (51).

Der Ausnahmefall (71) liefert in Verbindung mit (60) $f(1,1)\leqq a$, in Verbindung mit (57) $f(1,1)\geqq a$, also

$$f(1,1)=a. \tag{75}$$

Aus (71) und (75) folgt aber, daß die Minimalform (18) vorliegt, was nach der Voraussetzung ausgeschlossen ist.

Es mögen also die Beziehungen (10), (51), (72) gelten. Sollte

$$f(1,2)\geqq f(-5,2) \tag{76}$$

nicht erfüllt sein, so wende man die Substitution

$$\begin{cases} x=-x'-2y', \\ y=\qquad\quad y' \end{cases} \tag{77}$$

mit der Determinante -1, die zu sich selbst invers ist, an. Man erhält

$$a'=a,\qquad c'=f(-2,1),\qquad f'(-2,1)=f(0,1)=c.$$

Jede der Beziehungen (51) und (72) geht durch (77) in die gestrichene andere über. Ferner folgt aus (77)

$$f'(1,2)=f(-5,2),\qquad f'(-5,2)=f(1,2);$$

ist also (76) nicht erfüllt, so ist die gestrichene Beziehung (76) erfüllt. Ich nehme also im folgenden an, es seien (10), (51), (72) und (76) erfüllt.

Durch die Identität

$$5f(1,2)+f(-5,2)-30a=24c \tag{78}$$

geht (51) über in

$$5f(1,2)+f(-5,2)\leqq 6a; \tag{79}$$

durch die Identität

$$f(1,2)+5f(-5,2)-30a=24f(-2,1) \tag{80}$$

geht (72) über in

$$f(1,2)+5f(-5,2)\leqq 6a. \tag{81}$$

Da aber (76) besteht, so ist (81) eine Folge von (79) oder (72) eine Folge von (51). Ferner folgt aus (76) und (79)

$$6f(-5,2)\leqq 5f(1,2)+f(-5,2)\leqq 6a,$$

also

$$f(-5,2) \leqq a. \tag{82}$$

Die Substitution

$$\left\{\begin{aligned} x &= 5x' - 2y', \\ y &= -2x' + y' \end{aligned}\right. \tag{83}$$

mit der Determinante 1 ergibt

(84) $a' = f(5, -2) = f(-5, 2)$, (85) $c' = f(-2, 1)$,

(86) $f'(1, 2) = f(1, 0) = a$, (87) $f'(-5, 2) = f(-29, 12)$.

Es gilt die Identität

$$f(-29, 12) = 35(f(-5, 2) - a) + f(1, 2). \tag{88}$$

Ferner folgt aus (79) und (82)

$$f(1, 2) - a \leqq -\frac{1}{5}(f(-5, 2) - a) \leqq -(f(-5, 2) - a). \tag{89}$$

Es ist nach (87), (86), (88), (89), (82)

$$\begin{aligned} f'(-5, 2) - f'(1, 2) &= f(-29, 12) - a \\ = 35(f(-5, 2) - a) + (f(1, 2) - a) &\leqq 34(f(-5, 2) - a) \leqq 0. \end{aligned}$$

Es gilt also die gestrichene Beziehung (76). Ferner ist nach (85), (72), (82), (84)

$$c' = f(-2, 1) \leqq -a \leqq -f(-5, 2) = -a',$$

also gilt die gestrichene Beziehung (51). Aus dieser folgt aber in Verbindung mit der gestrichenen Beziehung (76) das Bestehen der gestrichenen Beziehung (72), wie mit Hilfe von (78) bis (81) gezeigt wurde. Solange nun

$$a' = f(-5, 2) > 0,$$

folgt, da wegen (82)

$$a' \leqq a,$$

aus (87), (88), (89), (84), (82)

$$\begin{aligned} \frac{f'(-5, 2) - a'}{a'} &= \frac{f(-29, 12) - f(-5, 2)}{a'} \\ = \left\{34\frac{f(-5, 2) - a}{a} + \frac{f(1, 2) - a}{a}\right\}\frac{a}{a'} &\leqq 33\frac{f(-5, 2) - a}{a} \leqq 0. \end{aligned}$$

Wenn also nicht zufällig

$$f(-5, 2) = a, \tag{90}$$

so tritt an die Stelle des negativen Ausdrucks $\frac{f(-5, 2) - a}{a}$ durch die Substitution (83) ein negativer, absolut mindestens 33 mal so großer Aus-

druck. Durch wiederholte Anwendung von (83) läßt sich also nach einer endlichen Anzahl von Schritten erreichen, daß

$$\frac{f^{(n)}(-5,2)-a^{(n)}}{a^{(n)}} \leqq -1 \quad \text{oder} \quad \frac{f^{(n)}(-5,2)}{a^{(n)}} \leqq 0.$$

Es ist also stets erreichbar, daß

$$\frac{f(-5,2)}{a} \leqq 0.$$

Sollte damit noch nicht erreicht sein, daß

$$(91) \qquad f(-5,2) \leqq -a,$$

wenn also die Bedingungen (10), (51) und

$$(92) \qquad -a < f(-5,2) < 0$$

gelten, so wende man die Substitution

$$(93) \qquad \begin{cases} x = -5x' - 13y', \\ y = 2x' + 5y' \end{cases}$$

mit der Determinante 1 an, und erhält nach (92)

$$\frac{f'(-5,2)}{a'} = \frac{f(-1,0)}{f(-5,2)} = \frac{a}{f(-5,2)} < -1.$$

Ferner folgt aus (93), (92), (82), (51)

$$\frac{c'}{a'} = \frac{f(-13,5)}{f(-5,2)} = \frac{\frac{13}{2}(f(-5,2)+a)-c}{f(-5,2)} < \frac{-c}{f(-5,2)} \leqq \frac{a}{f(-5,2)} < -1.$$

Nach Multiplikation mit -1 bestehen also die gestrichenen Beziehungen (10), (91), (51).

Der Ausnahmefall (90) liefert in Verbindung mit (79) $f(1,2) \leqq a$, in Verbindung mit (76) $f(1,2) \geqq a$, also

$$(94) \qquad f(1,2) = a.$$

Aus (90) und (94) folgt aber, daß die Minimalform (21) vorliegt, was nach der Voraussetzung ausgeschlossen ist.

Es mögen also die Beziehungen (10), (51) und (91) gelten. Es ist aber die Beziehung (51) identisch mit (5)(N), (91) identisch mit (20)(N). Es entstand aber (5)(N) aus (24)(N), (20)(N) aus (25)(N) durch Einsetzen der speziellen Werte (43), (44), (45).

Der Beweis wird von nun an durch Induktion geführt. Es genüge also $f(x,y)$ bereits (10) und einem Paare von Bedingungen

$$(24)(\mathrm{N}) \quad f(q_1, p_1) \leqq -a, \qquad (25)(\mathrm{N}) \quad f(q_2 - 3p_2, p_2) \leqq -a.$$

Nach Frobenius (F. M., § 3, S. 465, (5)) führe ich die Substitution ein

$$S_k = \begin{pmatrix} q_k, & -r_k \\ p_k, & -q_k \end{pmatrix},$$

wo der Index k gleich 1 oder 2 gesetzt oder fortgelassen werden kann; ferner

$$T_\varepsilon = \begin{pmatrix} \varepsilon, & 3 \\ 0, & \varepsilon \end{pmatrix},$$

worin also hier (23) gilt. Es bezeichne a_{S_k} den ersten Koeffizienten, der durch S_k transformierten Form $f(x, y)$. Es ist $a_{S_k} = a_{-S_k}$.

Es ist ferner

$$-T_{-1} S_2 T_1 = \begin{pmatrix} q_2 - 3p_2, & -9p_2 + 6q_2 - r_2 \\ p_2\;, & 3p_2 - q_2 \end{pmatrix},$$

$$S T_1 = \begin{pmatrix} q, & 3q - r \\ p, & 3p - q \end{pmatrix},$$

$$-(S T_1)^{-1} = \begin{pmatrix} q - 3p, & 3q - r \\ p\;, & -q \end{pmatrix},$$

$$(S T_1)^2 = \begin{pmatrix} 3pq - 1, & 3p(3q - r) \\ 3p^2\;, & 9p^2 - 3pq - 1 \end{pmatrix}.$$

Man erhält

$$a_{S_1} = f(q_1, p_1), \tag{95}$$

$$a_{-T_{-1}S_2T_1} = f(q_2 - 3p_2, p_2), \tag{96}$$

$$a_{ST_1} = f(q, p), \tag{97}$$

$$a_{-(ST_1)^{-1}} = f(q - 3p, p), \quad a_{(ST_1)^2} = f(3pq - 1, 3p^2). \tag{98}$$

Durch Elimination von b und c aus drei von diesen Gleichungen, unter Benutzung der Formel

$$\begin{vmatrix} x_1^2, & x_1 y_1, & y_1^2 \\ x_2^2, & x_2 y_2, & y_2^2 \\ x_3^2, & x_3 y_3, & y_3^2 \end{vmatrix} = \begin{vmatrix} x_1, & y_1 \\ x_3, & y_3 \end{vmatrix} \cdot \begin{vmatrix} x_2, & y_2 \\ x_3, & y_3 \end{vmatrix} \cdot \begin{vmatrix} x_1, & y_1 \\ x_2, & y_2 \end{vmatrix}$$

zur Ausrechnung der Determinanten dritten Grades, erhält man folgende Identitäten, die sich auch direkt verifizieren lassen:

$$\begin{aligned} a_{ST_1} \cdot p_1(p^2 + p_1^2) + a_{-(ST_1)^{-1}} \cdot p_1 p_2^2 - a \cdot 3 p p_2 (p^2 + p_1^2) \\ - a_{S_1} \cdot 3 p^3 p_2 = 0, \end{aligned} \tag{99}$$

$$\begin{aligned} a_{ST_1} \cdot p_1^2 p_2 + a_{-(ST_1)^{-1}} \cdot p_2 (p^2 + p_2^2) - a \cdot 3 p p_1 (p^2 + p_2^2) \\ - a_{-T_{-1}S_2T_1} \cdot 3 p^3 p_1 = 0, \end{aligned} \tag{100}$$

(101) $(a + a_{S_1})(p^2 + p_2^2) - (a_{-T_{-1}S_2T_1} + a_{ST_1})p_1^2 = 0,$

(102) $(a + a_{-T_{-1}S_2T_1})(p^2 + p_1^2) - (a_{S_1} + a_{-(ST_1)^{-1}})p_2^2 = 0,$

(103) $(a_{ST_1} - a)(9p^2 - 1) + a_{-(ST_1)^{-1}} - a_{(ST_1)^2} = 0.$

Wendet man (103) auf $a_{-(ST_1)^{-1}}$ an, so erhält man

(104) $(a_{-(ST_1)^{-1}} - a)(9p^2 - 1) + a_{ST_1} - a_{(ST_1)^{-2}} = 0.$

Die Bedingungen (24)(N) und (25)(N) nehmen nach (95) bzw. (96) die Gestalt an

(105) $a_{S_1} \leqq -a,$ (106) $a_{-T_{-1}S_2T_1} \leqq -a.$

Durch die Identität (99) geht (105) über in

(107) $a_{ST_1}(p^2 + p_1^2) + a_{-(ST_1)^{-1}} \cdot p_2^2 \leqq a(p^2 + p_1^2 + p_2^2);$

durch die Identität (100) geht (106) über in

(108) $a_{ST_1} \cdot p_1^2 + a_{-(ST_1)^{-1}}(p^2 + p_2^2) \leqq a(p^2 + p_1^2 + p_2^2).$

Man unterscheide zwei Hauptfälle, je nachdem, welche der beiden Zahlen a_{ST_1} und $a_{-(ST_1)^{-1}}$ die größere ist. Den Fall der Gleichheit ordne man nach Belieben einem der beiden Hauptfälle unter.

Hauptfall A: Es sei

(109) $a_{-(ST_1)^{-1}} \geqq a_{ST_1};$

dann ist (107) eine Folge von (108) oder (105) eine Folge von (106). Ferner folgt aus (109) und (108)

$$a_{ST_1}(p^2 + p_1^2 + p_2^2) \leqq a_{ST_1} \cdot p_1^2 + a_{-(ST_1)^{-1}}(p^2 + p_2^2) \leqq a(p^2 + p_1^2 + p_2^2)$$

oder

(110) $a_{ST_1} \leqq a.$

Aus (108) und (110) folgt ferner

(111) $a_{-(ST_1)^{-1}} - a \leqq -\frac{p_1^2}{p^2 + p_2^2}(a_{ST_1} - a) \leqq -(a_{ST_1} - a).$

Man mache die Substitution ST_1. Es soll gezeigt werden, daß die Beziehung (109) erhalten bleibt, daß also

(112) $a_{(ST_1)^2} - a \leqq 0.$

Es ist nach (103), (111), (110)

$$\begin{aligned} a_{(ST_1)^2} - a &= (a_{ST_1} - a)(9p^2 - 1) + (a_{-(ST_1)^{-1}} - a) \\ &\leqq (a_{ST_1} - a)(9p^2 - 2) \leqq 0, \end{aligned}$$

also besteht (112) oder die transformierte Beziehung (109). Es soll

ferner das Bestehen der transformierten Beziehung (106) nachgewiesen werden, d. h. es soll gezeigt werden, daß

(113) $$a_{-ST_1T_{-1}S_2T_1} \leqq - a_{ST_1}$$

oder unter Benutzung von F. M., § 3, S. 465 (6), daß

(114) $$a_{S_1} \leqq - a_{ST_1}.$$

Es ist aber nach (105) und (110)

$$a_{S_1} \leqq - a \leqq - a_{ST_1}.$$

Also ist (114) oder (113), d. h. die transformierte Beziehung (106) erfüllt. Aus dieser folgt aber in Verbindung mit der transformierten Beziehung (109) das Bestehen der transformierten Beziehung (105), wie mit Hilfe von (107) und (108) gezeigt wurde. Wenn nun

$$a_{ST_1} > 0$$

sein sollte, so folgt aus (103), (111), (110)

$$\frac{a_{(ST_1)^2} - a_{ST_1}}{a_{ST_1}} = \left\{(9p^2 - 2)\frac{a_{ST_1} - a}{a} + \frac{a_{-(ST_1)^{-1}} - a}{a}\right\}\frac{a}{a_{ST_1}}$$

$$\leqq (9p^2 - 3)\frac{a_{ST_1} - a}{a} \leqq 0.$$

Wenn also nicht zufällig

(115) $$a_{ST_1} = a,$$

so tritt an die Stelle des negativen Ausdrucks $\frac{a_{ST_1} - a}{a}$ durch die Substitution ST_1 ein negativer, absolut mindestens $9p^2 - 3$-mal so großer Ausdruck. Durch wiederholte Anwendung von ST_1 läßt sich also nach einer endlichen Anzahl von Schritten erreichen, daß

$$\frac{a_{(ST_1)^n} - a_{(ST_1)^{n-1}}}{a_{(ST_1)^{n-1}}} \leqq -1, \quad \text{oder daß} \quad \frac{a_{(ST_1)^n}}{a_{(ST_1)^{n-1}}} \leqq 0.$$

Es ist also stets erreichbar, daß $\frac{a_{ST_1}}{a} \leqq 0$. Sollte damit noch nicht erreicht sein, daß

(116) $$a_{ST_1} \leqq - a,$$

wenn also die Bedingungen (10), (106) und

(117) $$- a < a_{ST_1} = a_S < 0$$

gelten, so wende man die Substitution S an und erhält nach (117)

$$\frac{a_{S^2}}{a_S} = \frac{a}{a_S} < -1.$$

Es ist ferner zu zeigen, daß durch die Substitution S die Beziehung (106) erhalten bleibt, daß also

(118) $$\frac{a_{ST_{-1}S_2T_1}}{a_S} \leqq -1.$$

Wie bereits auseinandergesetzt wurde, darf man in den Formeln

$$p,\ p_1,\ p_2 \quad \text{durch} \quad p_1',\ p,\ p_2$$

ersetzen. Dabei gehen

$$S,\ S_1,\ S_2 \quad \text{über in} \quad S_1',\ S,\ S_2;$$

die Formel F. M., § 3, S. 465, (6.) geht über in

$$S_1' = S T_{-1} S_2 = S_2 T_{-1} S,$$

also geht (118) über in

(119) $$\frac{a_{S_1'T_1}}{a_S} \leqq -1$$

und die Identität (101) geht über in

$$(a + a_S)(p_1'^2 + p_2^2) - (a_{-T_{-1}S_2T_1} + a_{S_1'T_1})p^2 = 0.$$

Hieraus folgt aber in Verbindung mit (117), (106) und wieder (117)

$$a_{S_1'T_1} = (a + a_S)\frac{p_1'^2 + p_2^2}{p^2} - a_{-T_{-1}S_2T_1} > -a_{-T_{-1}S_2T_1} \geqq a > -a_S > 0,$$

als ist (119) oder (118), d. h. die transformierte Beziehung (106) erfüllt. Es läßt sich also nach Multiplikation mit -1 im Hauptfalle A. stets erreichen, daß (10), (106) und (116) gelten. Das sind aber wegen (96) und (97) die Beziehungen (27)(N) und (25)(N).

Der Ausnahmefalll (115) liefert in Verbindung mit (108) $a_{-(ST_1)^{-1}} \leqq a$, in Verbindung mit (109) $a_{-(ST_1)^{-1}} \geqq a$, also

(120) $$a_{-(ST_1)^{-1}} = a.$$

Den Gleichungen (115) und (120) genügt aber nur eine wohlbestimmte Form, und zwar mit Rücksicht auf (34) und (35) die Form (1) oder eine proportionale Form, die nach der Voraussetzung ausgeschlossen ist.

Völlig analog verläuft

Hauptfall B: Es sei

(121) $$a_{ST_1} \geqq a_{-(ST_1)^{-1}}.$$

Dann ist (108) eine Folge von (107) oder (106) eine Folge von (105). Ferner folgt aus (121) und (107)

$$a_{-(ST_1)^{-1}}(p^2 + p_1^2 + p_2^2) \leqq a_{ST_1}(p^2 + p_1^2) + a_{-(ST_1)^{-1}}p_2^2$$
$$\leqq a(p^2 + p_1^2 + p_2^2)$$

oder

$$(122) \qquad a_{-(ST_1)^{-1}} \leqq a.$$

Aus (107) und (122) folgt ferner

$$(123) \qquad a_{ST_1} - a \leqq - \frac{p_1^2}{p^2+p_1^2}(a_{-(ST_1)^{-1}} - a) \leqq - (a_{-(ST_1)^{-1}} - a).$$

Man mache die Substitution $-(ST_1)^{-1}$. Es soll gezeigt werden, daß die Beziehung (121) erhalten bleibt, daß also

$$(124) \qquad a_{(ST_1)^{-2}} - a \leqq 0.$$

Es ist nach (104), (123), (122)

$$\begin{aligned} a_{(ST_1)^{-2}} - a &= (a_{-(ST_1)^{-1}} - a)(9p^2-1) + (a_{ST_1} - a) \\ &\leqq (a_{-(ST_1)^{-1}} - a)(9p^2 - 2) \leqq 0, \end{aligned}$$

also besteht (124) oder die transformierte Beziehung (121). Es soll ferner das Bestehen der transformierten Beziehung (105) nachgewiesen werden, d. h. es soll gezeigt werden, daß

$$(125) \qquad a_{-(ST_1)^{-1}S_1} \leqq - a_{-(ST_1)^{-1}}$$

oder mit Benutzung von F. M., § 3, S. 465 (6), daß

$$(126) \qquad a_{-T_{-1}S_2T_1} \leqq - a_{-(ST_1)^{-1}}.$$

Es ist aber nach (106) und (122)

$$a_{-T_{-1}S_2T_1} \leqq - a \leqq - a_{-(ST_1)^{-1}}.$$

Also ist (126) und (125), d. h. die transformierte Beziehung (105) erfüllt. Aus dieser folgt aber in Verbindung mit der transformierten Beziehung (121) das Bestehen der transformierten Beziehung (106), wie mit Hilfe von (107) und (108) gezeigt wurde. Wenn nun

$$a_{-(ST_1)^{-1}} > 0$$

sein sollte, so folgt aus (104), (123), (122)

$$\begin{aligned} \frac{a_{(ST_1)^{-2}} - a_{-(ST_1)^{-1}}}{a_{-(ST_1)^{-1}}} &= \left\{(9p^2-2)\frac{a_{-(ST_1)^{-1}} - a}{a} + \frac{a_{ST_1} - a}{a}\right\}\frac{a}{a_{-(ST_1)^{-1}}} \\ &\leqq (9p^2 - 3)\frac{a_{-(ST_1)^{-1}} - a}{a} \leqq 0. \end{aligned}$$

Wenn also nicht zufällig

$$(127) \qquad a_{-(ST_1)^{-1}} = a,$$

so tritt an die Stelle des negativen Ausdrucks $\frac{a_{-(ST_1)^{-1}} - a}{a}$ durch die Substitution $-(ST_1)^{-1}$ ein negativer, absolut mindestens $9p^2 - 3$-mal so

großer Ausdruck. Durch wiederholte Anwendung von $-(S T_1)^{-1}$ läßt sich also nach einer endlichen Anzahl von Schritten erreichen, daß

$$\frac{a_{(ST_1)^{-n}} - a_{(ST_1)^{-(n-1)}}}{a_{(ST_1)^{-(n-1)}}} \leqq -1, \quad \text{oder daß} \quad \frac{a_{(ST_1)^{-n}}}{a_{(ST_1)^{-(n-1)}}} \leqq 0.$$

Es ist also stets erreichbar, daß $\frac{a_{-(ST_1)^{-1}}}{a} \leqq 0$. Sollte damit noch nicht erreicht sein, daß

$$(128) \qquad a_{-(ST_1)^{-1}} \leqq -a,$$

wenn also die Bedingungen (10), (105) und

$$(129) \qquad -a < a_{-(ST_1)^{-1}} = a_{-T_{-1}ST_1} < 0$$

gelten, so wende man die Substitution $-T_{-1}ST_1$ an und erhält nach (129)

$$\frac{a_{(T_{-1}ST_1)^2}}{a_{-T_{-1}ST_1}} = \frac{a}{a_{-T_{-1}ST_1}} \leqq -1.$$

Es ist ferner zu zeigen, daß durch die Substitution $-T_{-1}ST_1$ die Beziehung (105) erhalten bleibt, daß also

$$(130) \qquad \frac{a_{-T_{-1}ST_1S_1}}{a_{-T_{-1}ST_1}} \leqq -1.$$

Ersetzt man in den Formeln

$$p,\ p_1,\ p_2 \quad \text{durch} \quad p_2',\ p_1,\ p$$

so gehen

$$S,\ S_1,\ S_2 \quad \text{über in} \quad S_2',\ S_1,\ S$$

die Formel F. M., § 3, S. 465 (6) geht über in

$$S_2' = S_1 T_{-1} S = S T_1 S_1;$$

also geht (130) über in

$$(131) \qquad \frac{a_{-T_{-1}S_2'}}{a_{-T_{-1}ST_1}} = \frac{a_{-(S_2'T_1)^{-1}}}{a_{-T_{-1}S_2T_1}} \leqq -1$$

und die Identität (102) geht über in

$$(a + a_{-T_{-1}S_2T_1})(p_2'^2 + p_1^2) - (a_{S_1} + a_{-(S_2'T_1)^{-1}})p^2 = 0.$$

Hieraus folgt aber in Verbindung mit (129), (105) und wieder (129)

$$a_{-(S_2'T_1)^{-1}} = (a + a_{-T_{-1}ST_1})\frac{p_2'^2 + p_1^2}{p^2} - a_{S_1} > -a_{S_1} \geqq a > -a_{-T_{-1}ST_1} > 0,$$

also ist (131) oder (130), d. h. die transformierte Beziehung (105) erfüllt. Es läßt sich also nach Multiplikation mit -1 im Hauptfalle B. stets erreichen, daß (10), (105) und (128) gelten. Das sind aber wegen (95) und (98) die Beziehungen (24) (N) und (28) (N).

Der Ausnahmefall (127) liefert in Verbindung mit (107) $a_{ST_1} \leqq a$, in Verbindung mit (121) $a_{ST_1} \geqq a$, also

$$a_{ST_1} = a. \tag{132}$$

Es stimmen aber (127) und (132) überein mit (120) und (115), also handelt es sich wieder um die Form (1) oder eine proportionale Form, die nach der Voraussetzung ausgeschlossen ist.

Es läßt sich also eine Form, die (24)(N) und (25)(N) genügt, im Hauptfalle A. transformieren in eine Form, die (27)(N) und (25)(N) genügt, im Hauptfalle B. in eine Form, die (24)(N) und (28)(N) genügt. Daraus läßt sich aber, wie im Beweise der Behauptung II, schließen, daß durch geeignete Transformationen (46) erfüllbar ist. Damit ist die Behauptung III bewiesen.

§ 3.

Es soll im Anschluß an § 1 eine Übersicht über diejenigen Minimalformen gegeben werden, für die (8) erfüllt ist oder

$$\frac{D}{a^2} = 9. \tag{133}$$

Hilfssatz. *Eine Häufungsform von Minimalformen ist eine Minimalform.*

Es seien (a_n, b_n, c_n) Minimalformen.

$$\lim_{n \to \infty} (a_n, b_n, c_n) = (a, b, c).$$

Angenommen, es sei (a, b, c) keine Minimalform, dann gibt es ein von $0, 0$ verschiedenes Wertepaar ganzer Zahlen x, y, für das

$$|ax^2 + bxy + cy^2| = |\bar{a}| < |a|.$$

Man bestimme N so groß, daß für alle $n \geqq N$

$$|a - a_n| < \frac{|a| - |\bar{a}|}{4}, \qquad |a - a_n|\,x^2 < \frac{|a| - |\bar{a}|}{4},$$

$$|b - b_n| \cdot |xy| < \frac{|a| - |\bar{a}|}{4}, \qquad |c - c_n|\,y^2 < \frac{|a| - |\bar{a}|}{4}.$$

Dann wird

$$\begin{aligned} &|a_n x^2 + b_n xy + c_n y^2| \\ \leqq\ &|ax^2 + bxy + cy^2| + |a - a_n|\,x^2 + |b - b_n| \cdot |xy| + |c - c_n|\,y^2 \\ <\ &|\bar{a}| + 3\,\frac{|a| - |\bar{a}|}{4} = |a| - \frac{a - |\bar{a}|}{4} \\ <\ &|a| - |a - a_n| \leqq |a_n|. \end{aligned}$$

Das widerspricht aber der Voraussetzung, nach der (a_n, b_n, c_n) eine Minimalform sein sollte. Es ist also (a, b, c) eine Minimalform.

Unter der Voraussetzung

$$(134) \qquad \frac{D}{a^2} \leqq 9$$

nimmt (27) (n) die Gestalt an

$$(135) \qquad (2aq+bp)^2 \leqq Dp^2 - 4a^2 \leqq a^2(9p^2-4)$$

oder

$$(136) \qquad b \leqq a\left(\sqrt{9-\frac{4}{p^2}} - 2\frac{q}{p}\right),$$

ebenso nimmt (28) (n) die Gestalt an

$$(137) \qquad (2a(q-3p)+bp)^2 \leqq Dp^2 - 4a^2 \leqq a^2(9p^2-4)$$

oder

$$(138) \qquad b \geqq a\left(-\sqrt{9-\frac{4}{p^2}} + 2\left(3-\frac{q}{p}\right)\right).$$

(136) und (138) können nicht gleichzeitig erfüllt sein, also, so lange (134) gilt, auch nicht (27) (N) und (28) (N). Unter den Minimalformen, die (24) (N), (25) (N) und (134) genügen, gibt es außer (31) keine, für die

$$(139) \qquad \sqrt{9-\frac{4}{p^2}} - 2\frac{q}{p} < \frac{b}{a} < -\sqrt{9-\frac{4}{p^2}} + 2\left(3-\frac{q}{p}\right).$$

An dieser Stelle will ich eine Bemerkung einschalten, von der ich im folgenden keinen Gebrauch machen werde: *Unter der Voraussetzung* (134) *ist, wenn* (27) (n) *gilt,* (28) (p) *erfüllt, wenn* (28) (n) *gilt,* (27) (p) *erfüllt.*

Es ist nach (27) (n) oder (136)

$$(140)\begin{cases} (2a(q-3p)+bp)^2 - Dp^2 \geqq a^2\{(6p-\sqrt{9p^2-4})^2 - 9p^2\} \\ \qquad = a^2(36p^2 - 4 - 4\sqrt{81p^4-36p^2}) \\ \qquad > a^2(36p^2 - 4 - 4\sqrt{81p^4-36p^2+4}) \\ \qquad = a^2(36p^2 - 4 - 4(9p^2-2)) = 4a^2, \end{cases}$$

also ist (28) (p) erfüllt. Es ist ferner nach (28) (n) oder (138) und (140)

$$(2aq+bp)^2 - Dp^2 \geqq a^2\{(6p-\sqrt{9p^2-4})^2 - 9p^2\} > 4a^2,$$

also ist (27) (p) erfüllt.

Für die Formen, die (24) (N), (25) (N) und (134) genügen, ist nach (136) und (138)

$$(141) \qquad -\sqrt{9-\frac{4}{p_1^2}} + 2\left(3-\frac{q_1}{p_1}\right) \leqq \frac{b}{a} \leqq \sqrt{9-\frac{4}{p_1^2}} - 2\frac{q_1}{p_1}.$$

Es ist aber nach F. M., § 2, S. 462 (4)

$$\Delta = \left\{\sqrt{9-\frac{4}{p_1^2}} - 2\frac{q_1}{p_1}\right\} - \left\{-\sqrt{9-\frac{4}{p_2^2}} + 2\left(3-\frac{q_2}{p_2}\right)\right\}$$
$$= \sqrt{9-\frac{4}{p_1^2}} + \sqrt{9-\frac{4}{p_2^2}} - \frac{2p}{p_1 p_2}.$$

Bezeichnet man die größere der beiden Zahlen p_1 und p_2 mit p_α, die kleinere mit p_β, berechnet man p als größere Wurzel der quadratischen Gleichung (2), so ergibt sich, da $p_\alpha \geqq 2$, $p_\beta \geqq 1$,

$$\Delta = \sqrt{9-\frac{4}{p_\alpha^2}} + \sqrt{9-\frac{4}{p_\beta^2}} - 3 - \sqrt{9-\frac{4}{p_\alpha^2}-\frac{4}{p_\beta^2}} \tag{142}$$
$$= \frac{4}{p_\alpha^2}\left\{\frac{1}{\sqrt{9-\frac{4}{p_\beta^2}} + \sqrt{9-\frac{4}{p_\alpha^2}-\frac{4}{p_\beta^2}}} - \frac{1}{3+\sqrt{9-\frac{4}{p_\alpha^2}}}\right\}$$
$$\leqq \frac{4}{p_\alpha^2}\left(\frac{1}{\sqrt{8}+2} - \frac{1}{6}\right) < \frac{1}{6p_\alpha^2}.$$

p_α wächst aber mit p über alle Grenzen. Hält man a fest, so unterscheiden sich die Formen, die den Bedingungen (24) (N), (25) (N) und (134) genügen, mit Rücksicht auf (40) und (142) beliebig wenig voneinander.

Satz IV. *Eine Form, die der Bedingung* (133) *genügt und deren b in keinem der unendlich vielen Intervalle* (139) *liegt, ist eine Minimalform.*

Durch eine Substitution von der Form

$$x = x' + ky', \quad y = \pm y',$$

die zur Folge hat $a' = a$; $b' = 2ka \pm b$, läßt sich erreichen, daß

$$2 \leqq \frac{b'}{a'} \leqq 3. \tag{143}$$

Setzt man in (139) zunächst die speziellen Werte (43), dann (44) ein, so liefert das Nichterfülltsein von (139) anstatt (143) die schärfere Bedingung

$$5-\sqrt{8} \leqq \frac{b}{a} \leqq \sqrt{5}. \tag{144}$$

Das ist aber die Bedingung, die durch Einsetzen der speziellen Werte (43) und (44) aus (141) entsteht. Da (139) nicht erfüllt ist, folgt aber aus dem Bestehen von (141) das Bestehen der entsprechenden Bedingung, in der entweder p_1 oder p_2 durch p ersetzt ist. Da man diesen Schluß beliebig oft wiederholen kann, gibt es mit Rücksicht auf (40) und (142) in beliebiger Nähe der gegebenen Form Minimalformen von der Form (31), also ist die gegebene Form nach dem Hilfssatze eine Minimalform. Aus

demselben Grunde gibt es in beliebiger Nähe der gegebenen Form andere Minimalformen, für die (133) erfüllt ist. Die Menge der Werte $\frac{b}{a}$, die in Verbindung mit (133) Minimalformen liefern, ist also in sich dicht, da jeder Wert Häufungswert ist; sie ist ferner abgeschlossen, da sie nach dem Hilfssatze alle ihre Häufungswerte enthält. Die Menge ist also perfekt. Sie ist ferner nirgends dicht, da jedes Intervall Teilintervalle enthält, die von Werten der Menge frei sind.

Frobenius ordnet (F. M., § 8, S. 473) jedem irreduziblen positiven rationalen Bruche $\frac{\alpha}{\alpha'}$ eine Markoffsche Zahl durch Rekursion folgendermaßen zu.

Es sei β, β' die kleinste positive Lösung der Gleichung

$$\alpha\beta' - \beta\alpha' = -1, \tag{145}$$

γ, γ' die kleinste positive Lösung der Gleichung

$$\alpha\gamma' - \gamma\alpha' = 1, \tag{146}$$

so daß also $\alpha = \beta + \gamma$, $\alpha' = \beta' + \gamma'$. Dann ist $p_{\alpha\alpha'} = 3 p_{\beta\beta'} p_{\gamma\gamma'} - p_{\delta\delta'}$, worin $\delta = |\beta - \gamma|$, $\delta' = |\beta' - \gamma'|$, oder es ist, wie Frobenius (F. M., § 8, S. 474) zeigt, $p_{\alpha\alpha'}$ die größere Wurzel der Gleichung

$$p^2_{\alpha\alpha'} + p^2_{\beta\beta'} + p^2_{\gamma\gamma'} = 3\, p_{\alpha\alpha'} p_{\beta\beta'} p_{\gamma\gamma'}\,.$$

Satz V. *Die arithmetische Ordnung der Zahlen* $\frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}}$ *ist dieselbe wie die der Zahlen* $\frac{\alpha}{\alpha'}$.

(Die arithmetische Ordnung der Zahlen $\frac{b}{a}$ in den Formen (31) $\frac{b}{a} = 3 - 2\frac{q_{\alpha\alpha'}}{p_{\alpha\alpha'}}$ ist dieselbe wie die der Zahlen $\frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}}$.)

Man halte auch hier (23) fest, setze für p, p_1, p_2 bzw. $p_{\alpha\alpha'}, p_{\beta\beta'}, p_{\gamma\gamma'}$ und erhält aus F. M., § 2, S. 462 (4)

$$\frac{p_{\beta\beta'}}{q_{\beta\beta'}} - \frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}} = \frac{p_{\gamma\gamma'}}{q_{\beta\beta'}\, q_{\alpha\alpha'}} > 0, \qquad \frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}} - \frac{p_{\gamma\gamma'}}{q_{\gamma\gamma'}} = \frac{p_{\beta\beta'}}{q_{\alpha\alpha'}\, q_{\gamma\gamma'}} > 0,$$

also

$$\frac{p_{\beta\beta'}}{q_{\beta\beta'}} > \frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}} > \frac{p_{\gamma\gamma'}}{q_{\gamma\gamma'}}\,.$$

Ferner folgt aus (145)

$$\frac{\beta}{\beta'} - \frac{\alpha}{\alpha'} = \frac{1}{\alpha'\beta'} > 0 \tag{147}$$

und aus (146)

$$\frac{\alpha}{\alpha'} - \frac{\gamma}{\gamma'} = \frac{1}{\alpha'\gamma'} > 0, \tag{148}$$

also

$$\frac{\beta}{\beta'} > \frac{\alpha}{\alpha'} > \frac{\gamma}{\gamma'}.$$

Die arithmetische Ordnung der $\frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}}$ und der $\frac{\alpha}{\alpha'}$ stimmt also zunächst überein für konjugierte Markoffsche Zahlen. Da aber anfangend mit den Brüchen $\frac{1}{0}$ und $\frac{0}{1}$ durch Zwischenschalten der Mediante $\frac{\beta+\gamma}{\beta'+\gamma'}$ zwischen die Brüche $\frac{\beta}{\beta'}$ und $\frac{\gamma}{\gamma'}$ sukzessive alle positiven rationalen irreduziblen Brüche entstehen, so werden die entsprechenden $\frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}}$ in derselben Ordnung zwischengeschaltet. Da die Gleichungen F. M., § 2, S. 462 (4) erfüllt sind für die speziellen Werte (43), (44), (45), die der Bedingung (47) genügen, so ist (47) allgemein erfüllt.

Gegeben sei eine positive irrationale Zahl λ. Beginnend mit den Brüchen $\frac{1}{0}$ und $\frac{0}{1}$ schließe man λ zwischen zwei Brüche $\frac{\beta}{\beta'} > \lambda > \frac{\gamma}{\gamma'}$ ein, ersetze dann den einen der beiden Brüche durch ihre Mediante, so daß λ eingeschlossen bleibt, und denke sich das Verfahren unendlich oft wiederholt. Von einer bestimmten Stelle, $\frac{[\lambda]+1}{1} > \lambda > \frac{[\lambda]}{1}$, an wachsen die Nenner ständig ($[\lambda]$ bedeutet die größte ganze Zahl unterhalb λ), die Differenz zweier benachbarter Brüche wird also nach (147) und (148) beliebig klein; man erhält mithin eine gegen λ konvergierende Zahlenfolge. Dieser entspricht eine Folge von Formen (31), die gegen eine Form konvergieren, die der Bedingung (133) genügt.

Benutzt man eine rationale Zahl $\frac{\alpha}{\alpha'}$ unendlich oft zur Mediantenbildung, so ist

$$\lim_{n\to\infty} \frac{\beta + n\alpha}{\beta' + n\alpha'} = \frac{\alpha}{\alpha'}, \tag{149}$$

$$\lim_{n\to\infty} \frac{n\alpha + \gamma}{n\alpha' + \gamma'} = \frac{\alpha}{\alpha'}. \tag{150}$$

Zu (149) gehört diejenige Form $\varphi_\infty = (a_\infty, b_\infty, c_\infty)$, die der Bedingung (133) und wegen (33) der Bedingung $\varphi_\infty(q_{\alpha\alpha'} - 3\,p_{\alpha\alpha'}, p_{\alpha\alpha'}) = -a_\infty$ genügt. Ebenso gehört zu (150) diejenige Form $\varphi'_\infty = (a'_\infty, b'_\infty, c'_\infty)$, die der Bedingung (133) und wegen (32) der Bedingung $\varphi'_\infty(q_{\alpha\alpha'}, p_{\alpha\alpha'}) = -a'_\infty$ genügt. Läßt man in (134) bis (138) das Gleichheitszeichen gelten, so erhält man

$$(151)\qquad \frac{b_\infty}{a_\infty} = -\sqrt{9 - \frac{4}{p_{\alpha\alpha'}^2}} + 2\left(3 - \frac{q_{\alpha\alpha'}}{p_{\alpha\alpha'}}\right),$$

$$(152)\qquad \frac{b'_\infty}{a'_\infty} = \sqrt{9 - \frac{4}{p_{\alpha\alpha'}^2}} - 2\frac{q_{\alpha\alpha'}}{p_{\alpha\alpha'}}.$$

Dadurch sind φ_∞ und φ'_∞ bis auf einen Faktor bestimmt.

Genügt umgekehrt eine Minimalform den Bedingungen (133) und (144), so genügt sie Bedingungen (141) mit wachsenden Markoffschen Zahlen. Den $\frac{p_{\alpha\alpha'}}{q_{\alpha\alpha'}}$ entsprechen Brüche $\frac{\alpha}{\alpha'}$, die wieder eine konvergente Zahlenfolge bilden und eine Zahl λ definieren, die rational oder irrational sein kann.

Es entspricht also jeder irrationalen Zahl *eine*, jeder rationalen Zahl *zwei* Minimalformen, die den Bedingungen (133) und (144) genügen; den Zahlen 0 und ∞ entspricht je eine Minimalform. Umgekehrt entspricht jeder Minimalform, die (133) und (144) genügt, *eine* rationale oder irrationale Zahl. Auch für diese Zuordnung gilt Satz V.

(Eingegangen am 8. 8. 1923.)

Algebraische Theorie der zerlegbaren Ringe.
(Algebraische Theorie der Ringe. III.)

Von

Wolfgang Krull in Freiburg i. Br.

Die vorliegende Arbeit soll eine Fortführung und Ergänzung der Abhandlungen „Algebraische Theorie der Ringe I und II“[1]) darstellen. Dort wurde für eine gewisse Klasse von Ringen, die im wesentlichen endliche oder unendliche Systeme hyperkomplexer Größen mit kommutativer Multiplikation sind, der Begriff der algebraischen und transzendenten Erweiterung nach dem Vorbild der Körpertheorie eingeführt. Mit Hilfe der Übertragung der körpertheoretischen Methoden gelangt man dann zu einer *Typisierung* der betrachteten Ringe, wenigstens der vollkommenen, die dadurch ausgezeichnet sind, daß ihnen ein vollkommener Körper zugeordnet werden kann.

Im folgenden soll nun untersucht werden, wie die Verhältnisse, insbesondere die Fragen der Typisierung, in einem gewissen ausgezeichneten Falle, nämlich bei den sogenannten zerlegbaren Ringen, liegen. Der Begriff des zerlegbaren Ringes wurde von Herrn Fraenkel in seiner Dissertation[2]) axiomatisch festgelegt. Die zerlegbaren Ringe sind dadurch gekennzeichnet, daß sich in ihnen die Null im wesentlichen eindeutig als Produkt von Potenzen von „Primteilern“ (der Null) darstellen läßt. Axiomatisch kann man sie, wie in § 1 der vorliegenden Arbeit gezeigt wird, auch charakterisieren als *„Ringe, in denen es nur Nullteiler und Einheiten gibt, und in denen außerdem jedes Ideal ein Hauptideal ist, also aus der Gesamtheit aller durch ein festes Element teilbaren Ringelemente*

---

[1]) Math. Annalen **88**, S. 80–122, bzw. Math. Annalen **91**, S. 1–46. Diese Arbeiten werden in Zukunft mit A. I und A. II zitiert.

[2]) „Über die Teiler der Null und Zerlegung von Ringen“. Journal f. Math. **145**, S. 139–176. Diese Arbeit wird im folgenden mit „F.“ zitiert. Ferner wollen wir unter „N.“ die Arbeit „Idealtheorie in Ringbereichen“, Math. Annalen **83**, S. 24–67 von E. Noether verstehen.

besteht". Nach den allgemeinen Ergebnissen von A. II bzw. nach den Resultaten der Fraenkelschen Dissertation läßt sich jeder zerlegbare Ring als Summe von endlich vielen speziellen Ringen darstellen, wobei ein spezieller Ring dadurch ausgezeichnet ist, daß in ihm sich jeder Nullteiler als Produkt einer eindeutig bestimmten Potenz eines ein für allemal festgelegten Primteilers mit einer Einheit darstellen läßt. Wir hatten bereits in A. II mit einer ausgezeichneten Klasse von speziellen zerlegbaren Ringen, nämlich mit den sogenannten „Grundringen" zu tun. Im folgenden beschäftigen wir uns genau wie in den vorangehenden Arbeiten vorwiegend mit Ringen des speziellen Typs.

Zunächst werden die Vereinfachungen dargelegt, die sich hinsichtlich der Theorie der algebraischen und transzendenten Erweiterungen bei den zerlegbaren Ringen gegenüber dem allgemeinen Fall ergeben. Die leicht zugängliche Natur der zerlegbaren Ringe zeigt sich u. a. darin, daß wir den Aufbau der Polynomideale einer Variablen besser übersehen können als im allgemeinen Fall, und daß sich insbesondere für jedes Polynomideal eine Normalbasis einfacher Art aufstellen läßt. Des ferneren lassen sich, wie in § 3 gezeigt wird, die regulären Erweiterungen eingehender charakterisieren als im allgemeinen Fall und zwar durch den wichtigen Satz:

Die regulären Erweiterungen eines zerlegbaren speziellen Ringes sind die einzigen Erweiterungen, die wieder zu einem speziellen zerlegbaren Ring mit dem gleichen charakteristischen Exponenten führen.

Die Bedeutung dieses Theorems zeigt sich u. a. bei der Typisierung der vollkommenen Ringe. Man kann z. B. mit Hilfe unseres Satzes unmittelbar einsehen, daß die vollkommenen Grundringe die einzigen vollkommenen speziellen zerlegbaren Ringe sind, die den gleichen charakteristischen Exponenten wie ihr Primring besitzen.

Von besonderer Wichtigkeit ist bei der Typisierung noch eine weitere Untersuchung, die nur bei zerlegbaren Ringen durchführbar ist. Im allgemeinen Fall mußten wir uns, um zu befriedigenden Ergebnissen zu gelangen, auf „reguläre" Erweiterungen beschränken, durch die dem gegebenen Ring nach einer in A. II verwandten Ausdrucksweise gleichsam nur reguläre Elemente zugefügt werden. *Bei zerlegbaren Ringen beschäftigen wir uns auch mit allgemeinen algebraischen Erweiterungen, durch die neue Nullteiler eingeführt werden, und zwar mit solchen, die einen zerlegbaren Ring wieder in einen zerlegbaren verwandeln.* Es gelingt, derartige Erweiterungen erschöpfend zu charakterisieren, und einen genauen Einblick in ihren Aufbau zu gewinnen.

Aus dem gewonnenen Resultat ergeben sich wichtige Folgerungen für die *Typisierung der zerlegbaren Ringe.* Zunächst gewinnen wir mit Hilfe

der allgemeinen algebraischen Erweiterungen einen befriedigenden Einblick in die Natur der *endlichen zerlegbaren Ringe*, und zwar zeigt es sich, daß man für alle hier in Betracht kommenden Typen Beispiele gewinnen kann, wenn man Restklassensysteme nach geeignet gewählten Idealen aus algebraischen Zahlkörpern betrachtet. Allerdings ist damit noch nicht bewiesen, daß jeder endliche zerlegbare Ring dem Restklassensystem eines algebraischen Ideales isomorph ist. Um diese Frage beantworten zu können, müßte man, wie im Texte dargelegt wird, jedenfalls erst über eine Reihe von Existenzsätzen aus der algebraischen Zahlentheorie Bescheid wissen.

Ähnlich wie bei den endlichen liegen die Verhältnisse bei beliebigen *vollkommenen zerlegbaren Ringen*. Auch hier gelangen wir auf Grund der Tatsache, daß *jeder vollkommene zerlegbare Ring aus seinem Grundring durch allgemeine algebraische Erweiterung entsteht*, zu einer genauen Charakterisierung der untersuchten Bereiche. *Besonders interessant sind die Verhältnisse bei den algebraisch abgeschlossenen zerlegbaren Ringen. Hier gelingt es uns nämlich, notwendige und hinreichende Kennzeichen dafür zu gewinnen, wann ein solcher Ring durch Grundring und charakteristischen Exponenten eindeutig bestimmt ist und wann nicht.* Dabei zeigt sich eine neue beachtenswerte Analogie mit der Körpertheorie: Der Unterschied zwischen den durch Grundring und charakteristischen Exponenten eindeutig bestimmten Ringen und den andern ist ganz ähnlich dem Unterschied zwischen vollkommenen und unvollkommenen Körpern (obwohl alle hier in Betracht kommenden Ringe in unserer in A. I und A. II eingeführten Ausdrucksweise vollkommen sind) — die Primteiler der Null spielen eine analoge Rolle wie die transzendenten Elemente in der Körpertheorie.

Außer diesen neuen Beziehungen zur Körpertheorie ergibt sich bei den zuletzt besprochenen Untersuchungen noch eine Reihe merkwürdiger Fragen, so daß es scheint, als ob die hier entwickelte algebraische Behandlungsweise nicht nur hinsichtlich der Typisierung der zerlegbaren Ringe zu einem gewissen Abschluß führte, sondern auch den geeigneten Ausgangspunkt zu einer weiteren Untersuchung dieser Bereiche böte.

§ 1.

Die allgemeinen zerlegbaren Ringe.

Ein allgemeiner zerlegbarer Ring ist ein Ringbereich mit kommutativer Multiplikation und universellem Einheitselement[3]) r_e, der folgenden Axiomen genügt:

[3]) Zu der Definition des allgemeinsten kommutativen Ringbereichs vgl. A. I § 1.

I. *Ein Element a aus R ist entweder Nullteiler oder Einheit, d. h. es besteht stets entweder eine Gleichung $a\cdot\bar a = 0$; $\bar a \neq 0$ oder eine Gleichung $a\cdot a^{-1} = r_e$.*

II. *Jedes Ideal aus R ist ein Hauptideal (a), besteht also aus der Gesamtheit der durch ein festes Element a teilbaren Ringelemente.*

Auf Grund dieser beiden Axiome beweisen wir zunächst folgenden Satz[4]):

Hilfssatz 1. *Besitzt das Ideal $\mathfrak{a} = (a)$ einen vom Einheitsideal $\mathfrak{o}$ verschiedenen echten Teiler $\mathfrak{b} = (b)$, so läßt es sich auch als Produkt echter Teiler darstellen.*

Da die Ideale $\mathfrak{a}$ und $\mathfrak{b}$ vom Einheitsideal verschieden sind, so müssen infolge von Axiom I die Elemente a und b Nullteiler sein, und wegen der Teilbarkeit von $\mathfrak{a}$ durch $\mathfrak{b}$ muß eine Gleichung $a = b\cdot c$ bestehen, aus der dann die Idealgleichung $(a) = (b)\cdot(c)$ folgt. Sollte nun (c) ein echter Teiler von $\mathfrak{a}$ sein, so ist unsere Behauptung bewiesen, im andern Falle muß jedenfalls eine Gleichung $c = d\cdot a$ bestehen. Ist dann $\bar b \neq 0$ ein nach Vor. existierendes Element, das der Gleichung $b\cdot\bar b = 0$ genügt, so kann $\bar b$ und folglich auch $\bar b + d\cdot a = \bar b + c$ nicht durch a teilbar sein, weil sonst entgegen der Wahl von $\bar b$ die Gleichung $\bar b = e\cdot a = e\cdot b\cdot c = e\cdot a\cdot b\cdot d = \bar b\cdot b\cdot d = 0$ bestände. Die Ideale $(\bar b + c)$ und $\mathfrak{a}$ sind mithin verschieden, und aus der Gleichung $\mathfrak{a} = (\bar b + c)\cdot(b)$ folgt die Behauptung des Hilfssatzes.

Da ein Primideal seiner Definition gemäß nicht als das Produkt von echten Teilern darstellbar ist, so folgt aus dem gewonnenen Ergebnis unmittelbar, daß in R kein Primideal einen von $\mathfrak{o}$ verschiedenen echten Teiler besitzen kann. Daraus ergibt sich weiter, daß *die zerlegbaren Ringe dem allgemeinen in* A. II § 1 *u.* § 2 *charakterisierten Ringtypus angehören,* und wir können aus A. II die folgenden Sätze übernehmen, die sich für zerlegbare Ringe mit Benutzung von Hilfssatz 1 auch direkt beweisen lassen:

Jeder allgemeine zerlegbare Ring läßt sich als eindeutige Summe von endlich viel speziellen Ringen, in denen die Gesamtheit der Nullteiler ein Primideal bildet, darstellen[5]).

Ist $\mathfrak{a}$ ein beliebiges Ideal aus R, so gilt eine eindeutige Produktzerlegung $\mathfrak{a} = \mathfrak{q}_1\cdot\mathfrak{q}_2\cdot\ldots\cdot\mathfrak{q}_\sigma$, wobei die $\mathfrak{q}_i$ gegenseitig teilerfremde Primärideale bedeuten[6]).

Zu jedem Ideal $\mathfrak{a}$ gibt es einen endlichen Exponenten $\varrho_\mathfrak{a}$, so daß $\mathfrak{a}^{\varrho_\mathfrak{a}} \neq \mathfrak{a}^{\varrho_\mathfrak{a}+1}$ wird, insbesondere existiert zu einem beliebigen speziellen zer-

---

[4]) Hilfssatz 1 ist im wesentlichen mit Satz 2, F. § 2 identisch.

[5]) A. II § 2, vgl. ferner die Entwicklungen bei F. § 4, wo der Zerlegungssatz ohne idealtheoretische Hilfsmittel bewiesen ist.

[6]) A. II. § 2.

legbaren Ring eine natürliche Zahl ϱ, *so daß für* $\mathfrak{a} \neq 0$ *immer* $\mathfrak{a}^\varrho = (0)$ *ist, oder, was dasselbe bedeutet, ein Produkt von* ϱ *Nullteilern stets verschwindet*[7]).

Aus der Tatsache, daß jedes Ideal aus R ein Hauptideal ist, ergibt sich ferner die wichtige Folgerung: *Es gibt in* R *keine ins Unendliche laufende Kette von Idealen* $\mathfrak{a}_1, \mathfrak{a}_2, \ldots, \mathfrak{a}_i, \ldots$, *bei der allgemein das Ideal* $\mathfrak{a}_i$ *ein echter Teiler von* $\mathfrak{a}_{i-1}$ *ist*[8]).

Denn der größte gemeinschaftliche Teiler der Ideale $\mathfrak{a}_i$ ist selbst ein Ideal $\mathfrak{d} = (d)$ und unter unsern Vor. muß das Element d für ein endliches ν im Ideale $\mathfrak{a}_\nu$ auftreten. Dann aber wird $\mathfrak{a}_\nu = (d)$, und die Kette bricht mit $\mathfrak{a}_\nu$ ab.

Aus dem zuletzt gewonnenen Ergebnis und aus Hilfssatz 1 folgt leicht:

Satz 1. *Für jedes Ideal* $\mathfrak{a}$ *aus* R *gilt eine Gleichung*

$$\mathfrak{a} = \mathfrak{p}_1^{e_1} \cdot \mathfrak{p}_2^{e_2} \cdot \ldots \cdot \mathfrak{p}_\sigma^{e_\sigma},$$

wobei die $\mathfrak{p}_i$ *teilerfremde Primideale bedeuten.*

Da in einem zerlegbaren Ring zwei verschiedene Primideale stets teilerfremd sind, so brauchen wir nur zu zeigen, daß sich jedes Ideal $\mathfrak{a}$ als Produkt von Primidealen darstellen läßt, dann ergibt sich durch Zusammenfassung gleicher Faktoren die Behauptung.

Sollte nun $\mathfrak{a}$ nicht als Produkt endlich vieler Primideale darstellbar sein, so müßte nach Hilfssatz 1 sicher eine Gleichung $\mathfrak{a} = \mathfrak{a}_1 \cdot \mathfrak{a}_2$ bestehen, wobei $\mathfrak{a}_1$ und $\mathfrak{a}_2$ echte Teiler von $\mathfrak{a}$ wären (denn jedes Ideal ohne von $\mathfrak{o}$ verschiedenen echten Teiler ist ja evidenterweise selbst Primideal). $\mathfrak{a}_1$ und $\mathfrak{a}_2$ können nun nicht beide als Produkt endlich vieler Primideale darstellbar sein, denn sonst gälte ja wegen $\mathfrak{a} = \mathfrak{a}_1 \cdot \mathfrak{a}_2$ das gleiche von $\mathfrak{a}$. Ist nun $\mathfrak{a}_1$ nicht als Produkt endlich vieler Primideale darstellbar, so haben wir eine Gleichung $\mathfrak{a}_1 = \mathfrak{a}_{11} \cdot \mathfrak{a}_{12}$, wobei, wie aus den vorhin bei $\mathfrak{a}$ angewandten Schlüssen folgt, bei geeigneter Bezeichnung $\mathfrak{a}_{11}$ einen nicht als Produkt von endlich viel Primidealen darstellbaren echten Teiler von $\mathfrak{a}_1$ bedeutet. Indem man jetzt auf $\mathfrak{a}_{11}$ dieselbe Schlußweise anwendet, wie früher auf $\mathfrak{a}$ und $\mathfrak{a}_1$ usw., erkennt man, daß jedes Ideal $\mathfrak{a}$ aus R als Produkt endlich vieler Primideale darstellbar sein muß, weil es andernfalls in R eine ins Unendliche laufende echte Teilerkette $\mathfrak{a}, \mathfrak{a}_1, \mathfrak{a}_{11}, \ldots$ gäbe.

Da in einem zerlegbaren Ring nach den in A. II gewonnenen allgemeinen Resultaten[9]) ein Ideal $\mathfrak{q}$ dann und nur dann Primärideal ist, wenn

---

[7]) A. II. § 2.

[8]) Vgl. N. § 1 S. 30, wo der entsprechende Satz unter der allgemeineren Vor. bewiesen ist, daß jedes Ideal aus R eine endliche Basis besitzt.

[9]) Vgl. A. II.

es durch ein einziges (von $\mathfrak{o}$ verschiedenes) Primideal teilbar ist, so können wir aus Satz 1 unmittelbar folgern:

Satz 2. *Die Potenzen der Primideale, und nur diese, sind in R Primärideale*[10]).

Die Produktzerlegung $\mathfrak{a} = \mathfrak{p}_1^{\varrho_1} \cdot \mathfrak{p}_2^{\varrho_2} \cdot \ldots \cdot \mathfrak{p}_\sigma^{\varrho_\sigma}$ *ist also mit der Produktzerlegung* $\mathfrak{a} = \mathfrak{q}_1 \cdot \mathfrak{q}_2 \cdot \ldots \cdot \mathfrak{q}_\sigma$, *die wir oben auf Grund der allgemeinen Ergebnisse von* A. II *eingeführt haben, identisch, d. h. es ist bei geeigneter Numerierung* $\mathfrak{p}_i^{\varrho_i} = \mathfrak{q}_i$ $(i = 1, 2, \ldots, \sigma)$. *Dabei sind die Exponenten* ϱ_i *eindeutig bestimmt, falls man sie stets so klein wählt, daß* $\mathfrak{p}_i^{\varrho_i - 1} \neq \mathfrak{p}_i^{\varrho_i}$ *wird.*

Aus Satz 1 *und* 2 *geht hervor, daß die durch Axiom* I *und* II *charakterisierten zerlegbaren Ringe mit den von Herrn Fraenkel behandelten „zerlegbaren Ringen mit endlich viel wesentlich verschiedenen Nullteilern" identisch sind.*

Sind a_1 und a_2 zwei Basiselemente desselben Ideals — wir wollen a_1 und a_2 in diesem Fall äquivalent nennen —, so unterscheiden sich a_1 und a_2 nach den allgemeinen Ergebnissen von A. II bzw. F. nur um einen Einheitsfaktor. Bezeichnen wir jedes Basiselement eines Primideals als „Primteiler" (der Null), so können wir daher sagen: *Die zerlegbaren Ringe sind diejenigen Ringbereiche, in denen es nur Nullteiler und Einheiten gibt, und in denen sich jedes Element bis auf Einheitsfaktoren eindeutig als Produkt von Primteilern darstellen läßt.* Von dieser Definition der zerlegbaren Ringe kann man ausgehen, falls man die Idealtheorie vermeiden will.

Die speziellen Ringe, auf die man einen allgemeinen zerlegbaren Ring zurückführen kann, sind, wie aus den Ergebnissen des zweiten Teils bzw. der Fraenkelschen Arbeit folgt, und wie man übrigens auch leicht mit Hilfe der Sätze 1 und 2 zeigen kann, ebenfalls zerlegbar, und *zwar bestehen die sämtlichen vom Einheitsideal verschiedenen Ideale eines speziellen zerlegbaren Rings aus dem Nullteilerprimideal* $\mathfrak{p}^* = (p)$ *und dessen Potenzen* $\mathfrak{p}^{*2} = (p^2), \ldots, \mathfrak{p}^{*\varrho} = (p^\varrho) = (0)$.[11]) Es ist somit bei den speziellen zerlegbaren Ringen eine noch viel weitergehende Analogie mit dem Restklassensystem nach einer Primzahlpotenz vorhanden, als wie wir sie bei dem im ersten Teil behandelten umfassenderen Fall fanden.

---

[10]) Man beachte, daß gleichzeitig mit $\mathfrak{p}$ auch jede Potenz von $\mathfrak{p}$ zu allen von $\mathfrak{p}$ verschiedenen Primidealen teilerfremd ist!

[11]) Vgl. den Hilfssatz von A. II § 2, sowie F. § 4 Satz 7. (Nach der bei F. gegebenen Definition sind ja die bei F. als „einfach", hier als „speziell" bezeichneten zerlegbaren Ringe im wesentlichen durch die eben formulierte Eigenschaft der Ideale charakterisiert.)

Die wichtigsten Beispiele für zerlegbare Ringe sind neben den Restklassensystemen nach einer natürlichen Zahl die Restklassensysteme nach einem Ideal eines algebraischen Zahlkörpers. Wir werden auf das letztere Beispiel in § 4 bei Behandlung der endlichen zerlegbaren Ringe zurückkommen.

§ 2.

Polynomideale in speziellen zerlegbaren Ringen.

Es sei R ein beliebiger spezieller zerlegbarer Ring. Mit $\mathfrak{p}^*$ bezeichnen wir gemäß der bereits in § 1 benutzten Terminologie sein Nullteilerprimideal, unter p wollen wir stets einen Primteiler, also ein Basiselement von $\mathfrak{p}^*$, verstehen. *Da nach einer in* § 1 *gemachten Bemerkung die Ideale* $\mathfrak{o}, \mathfrak{p}^*, \mathfrak{p}^{*2}, \ldots, \mathfrak{p}^{*\varrho} = (0)$ *die einzigen in R vorkommenden Ideale sind, so läßt sich jedes Element aus R in der Form $a = r \cdot p^{\sigma}$ darstellen, wobei r eine Einheit bedeutet. Dabei ist der Exponent σ für $a \neq 0$ eindeutig bestimmt.*

Die Theorie der algebraischen und transzendenten Erweiterungen eines speziellen zerlegbaren Ringes läßt sich nun z. T. wesentlich einfacher gestalten, z. T. beträchtlich weiter ausbauen, als es im allgemeinen möglich gewesen war, und zwar ist einer der Gründe der, daß, wie in diesem Paragraphen gezeigt werden soll, in unserm Sonderfall ein tieferer Einblick in die Natur des zu R gehörigen Polynomrings R_f möglich ist.

Es sollen zunächst die Ideale aus R_f untersucht werden. Unter $\mathfrak{p}_f^*$ wollen wir in üblicher Weise das aus allen Nullteilern bestehende Primideal verstehen. Die Ideale $\mathfrak{p}_f^{*i}$ $(i = 1, 2, \ldots, \varrho)$ sind dann (ebenso wie die Ideale $\mathfrak{p}^{*i}$ aus R) Hauptideale, denn es ist $\mathfrak{p}_f^{*i} = (p^i)$. Es sei jetzt $\mathfrak{a}$ ein beliebiges reguläres oder nicht reguläres Ideal aus R_f. Wir betrachten die Reihe der Ideale $\mathfrak{a} = \mathfrak{a}_0 = \mathfrak{a} : \mathfrak{p}^{*0}$; $\mathfrak{a}_1 = \mathfrak{a} : \mathfrak{p}^*$; $\mathfrak{a}_2 = \mathfrak{a} : \mathfrak{p}^{*2}$ usw. Nach der Definition des Idealquotienten[12]) ist $\mathfrak{a}_i$ der größte gemeinschaftliche Teiler aller derjenigen Ideale $\mathfrak{b}$, für die die Kongruenz $\mathfrak{b} \cdot \mathfrak{p}^{*i} \equiv 0 (\mathfrak{a})$ gilt, und es stellt daher $\mathfrak{a}_{i+1}$ einen (echten oder unechten) Teiler von $\mathfrak{a}_i$ dar. Wegen $\mathfrak{a} : \mathfrak{p}^{*\varrho} = \mathfrak{a} : (0) = (r_e)$ kommt in der Reihe $\mathfrak{a}_0, \mathfrak{a}_1, \ldots, \mathfrak{a}_\varrho$ sicher ein reguläres Ideal vor.

Satz 3. *Ist λ die kleinste natürliche Zahl, für die $\mathfrak{a}_\lambda$ regulär ist, so gilt die Gleichung* $\mathfrak{a} = \mathfrak{a}_\lambda \cdot \mathfrak{p}^{*\lambda}$.

Es ist wegen $\mathfrak{a}_\lambda \cdot \mathfrak{p}^{*\lambda} \equiv 0 (\mathfrak{a})$ nur nachzuweisen, daß auch umgekehrt die Kongruenz $\mathfrak{a} \equiv 0 (\mathfrak{a}_\lambda \cdot \mathfrak{p}^{*\lambda})$ gilt. Das ist aber unter unsern Voraussetzungen über λ trivial, denn es muß jedes Element aus $\mathfrak{a}$ die Gestalt

---

[12]) Vgl. A. I § 1.

$a(x) = p^\lambda \cdot a'(x)$ besitzen, wobei p^λ durch $\mathfrak{p}^{*\lambda}$ und $a'(x)$ durch $\mathfrak{a}_\lambda$ teilbar ist. Satz 3 zeigt, daß sich in R_f die Theorie der nicht regulären Ideale auf die der regulären zurückführen läßt. Wir beschäftigen uns daher in Zukunft im wesentlichen mit den regulären Idealen.

Satz 4. *Jedes reguläre*[13]) *Ideal aus R_f besitzt eine endliche Basis, und zwar von höchstens ϱ Elementen.*

Bedeutet nämlich $f_i(x)$ eine reguläre Funktion niedrigsten Grades aus $\mathfrak{a}_i$ $(i = 0, 1, \ldots, \varrho - 1)$, so gilt die Gleichung

$$(f_0(x), p \cdot f_1(x), \ldots, p^{\varrho-1} \cdot f_{\varrho-1}(x)) = \mathfrak{a}.$$

In der Tat, ist $a(x)$ ein beliebiges Element aus $\mathfrak{a}$, so haben wir

$$a(x) = b_0(x) \cdot f_0(x) + p \cdot a_1(x); \qquad a_1(x) \equiv 0(\mathfrak{a}_1),$$

also

$$a(x) = b_0(x) \cdot f_0(x) + p \cdot b_1(x) \cdot f_1(x) + p^2 \cdot a_2(x); \quad a_2(x) \equiv 0(\mathfrak{a}_2); \ \ldots;$$

$$a(x) = b_0(x) \cdot f_0(x) + \ldots + p^{\varrho-1} \cdot b_{\varrho-1}(x) \cdot f_{\varrho-1}(x) + 0 \cdot a_\varrho(x)$$
$$= \sum_{i=0}^{\varrho-1} b_i(x) \cdot f_i(x) \cdot p^i.$$

Eine Basis der eben angegebenen Art, bei der das $i + 1$-te Glied das Produkt von p^i mit einer regulären Funktion niedrigsten Grades aus $\mathfrak{a}_i$ ist, soll als *Normalbasis von* $\mathfrak{a}$ bezeichnet werden. Besitzen nun die Funktionen niedrigsten Grades aus $\mathfrak{a}_i$ und $\mathfrak{a}_{i+1}$ den gleichen Grad, so ist jede Funktion niedrigsten Grades aus $\mathfrak{a}_i$ auch eine solche aus $\mathfrak{a}_{i+1}$, und daraus folgt unmittelbar, daß unter diesen Umständen das $i + 2$-te Glied in einer Normalbasis von $\mathfrak{a}$ einfach weggelassen werden kann. Auf Grund dieser Tatsache können wir zu jedem Ideal $\mathfrak{a}$ eine „*gekürzte Normalbasis*"

$$\mathfrak{a} = (f_0(x), p^{\mu_1} \cdot f_1(x), p^{\mu_1+\mu_2} \cdot f_2(x), \ldots, p^{\mu_1+\mu_2+\cdots+\mu_\sigma} \cdot f_\sigma(x))$$

konstruieren, bei der $f_i(x)$ eine Funktion niedrigsten Grades aus $\mathfrak{a}_{\mu_1+\mu_2+\cdots+\mu_i}$ bedeutet, und bei der insbesondere der Grad von $f_i(x)$ stets niedriger ist, als der von $f_{i-1}(x)$ $(i = 1, 2, \ldots, \sigma)$. Wir bemerken noch:

Ist $g(x)$ Primfunktion, und ist $\mathfrak{a}$ ein zum Primideal

$$\mathfrak{p} = ((g(x)), \mathfrak{p}_f^*) = (g(x), p) \text{ [14])}$$

gehöriges Primärideal, bedeutet ferner $(f_0(x), p \cdot f_1(x), p^2 \cdot f_2(x), \ldots)$ eine Normalbasis von $\mathfrak{a}$, so sind die Funktionen $f_i(x)$ sämtlich zu $\mathfrak{p}$ gehörige Primärfunktionen oder Einheiten.

---

[13]) Also wegen Satz 3 und wegen der Gleichung $\mathfrak{p}^{*\lambda} = (p^\lambda)$ auch jedes nichtreguläre.

[14]) Vgl. A. I § 6 Satz 11.

In der Tat, zunächst ist nach A. I § 6 Satz 12 jedenfalls $f_0(x)$ eine zu $\mathfrak{p}$ gehörige Primärfunktion. Es sei nun (für $i \geqq 1$) $f_i(x) = h_i(x) \cdot \bar{h}_i(x)$, wobei $h_i(x)$ eine zu $\mathfrak{p}$ gehörige Primärfunktion oder eine Einheit bedeutet, während $\bar{h}_i(x)$ zu $\mathfrak{p}$ teilerfremd ist. Dann ist neben $p^i \cdot h_i(x) \cdot \bar{h}_i(x)$ auch $p^i \cdot f_0(x)$ durch $\mathfrak{a}$ teilbar, und daraus ergeben sich wegen der Teilerfremdheit von $f_0(x)$ und $\bar{h}_i(x)$ die Kongruenzen $p^i \cdot h_i(x) \equiv 0(\mathfrak{a})$, $h_i(x) \equiv 0(\mathfrak{a}_i)$. Da nun $f_i(x)$ eine Funktion niedrigsten Grades aus $\mathfrak{a}_i$ bedeutet, so muß $h_i(x)$ eine Einheit sein, d. h. $f_i(x)$ ist, wie behauptet, primär, und gehört zu $\mathfrak{p}$, oder es ist selbst Einheit. Wir können in Anbetracht von A. I § 6 Satz 12 das gewonnene Resultat auch so formulieren:

Ist $\mathfrak{a}$ primär, so sind auch die Ideale $\mathfrak{a}_i = \mathfrak{a} : \mathfrak{p}^{*i}$, soweit sie vom Einheitsideal verschieden sind, sämtlich primär und gehören zum selben Primideal wie $\mathfrak{a}$.[15])

Für das Folgende wollen wir noch eine kurze Bemerkung machen.

Es sei $\mathfrak{a}$ ein zum Primideal $\mathfrak{p} = (g(x), p)$ gehöriges Primärideal, $(g(x)^\mu + q(x), p \cdot f_1(x) \ldots)$ sei eine Normalbasis von $\mathfrak{a}$, und es sei $\mu > 1$, es möge also $\mathfrak{a}$ keine Primfunktion enthalten. Wir fragen uns, wann dann $\mathfrak{a}$ durch $\mathfrak{p}^2$ teilbar ist.

Soll $\mathfrak{a}$ durch $\mathfrak{p}^2$ teilbar sein, so darf $f_1(x)$ keine Einheit sein, es muß also die Kongruenz $f_1(x) \equiv 0(\mathfrak{p})$ bestehen. Ist diese Bedingung erfüllt, so sind $p \cdot f_1(x), p^2 \cdot f_2(x), \ldots$ sämtlich durch $\mathfrak{p}^2$ teilbar, es dreht sich also nur noch um die Frage, wann $g(x)^\mu + q(x)$ in $\mathfrak{p}^2$ enthalten ist. Soll die Kongruenz $g(x)^\mu + q(x) \equiv 0(\mathfrak{p}^2)$ bestehen, so muß um so mehr $g(x)^\mu + q(x) \equiv 0\big((g(x), p^2)\big)$ sein. Aus der letzteren Kongruenz folgt, daß $q(x)$ die Gestalt $q(x) = p \cdot g(x) \cdot h(x) + p^2 \cdot h_1(x)$ besitzen, also durch $\mathfrak{p}^2 = (g(x)^2, p \cdot g(x), p^2)$ teilbar sein muß. Wegen $\mu > 1$ ist dann auch $g(x)^\mu$ durch $\mathfrak{p}^2$ teilbar, es folgt also für $\mu > 1$ aus der Kongruenz $g(x)^\mu + q(x) \equiv 0\big((g(x), p^2)\big)$ rückwärts die Kongruenz $g(x)^\mu + q(x) \equiv 0(\mathfrak{p}^2)$ und daraus weiter $\mathfrak{a} \equiv 0(\mathfrak{p}^2)$.

Das gewonnene Resultat formulieren wir mit Rücksicht auf das Folgende etwas spezieller als nötig so:

Es sei $\mathfrak{a}$ ein zum regulären Primideal $\mathfrak{p}$ gehöriges Primärideal, das weder eine Primfunktion noch ein Element aus R, also insbesondere keinen Primteiler aus R enthält, $f(x)$ sei ein reguläres Polynom niedrigsten Grades aus $\mathfrak{a}$, $g(x)$ sei eine beliebige Primfunktion aus $\mathfrak{p}$. Dann ist die notwendige und hinreichende Bedingung für die Teilbarkeit von $\mathfrak{a}$ durch $\mathfrak{p}^2$ das Bestehen der Kongruenz $f(x) \equiv 0\big((g(x), p^2)\big)$.

---

[15]) Dies Ergebnis kann auch als Spezialfall aus Satz 3 der Note von W. Krull: Ein neuer Beweis für die Hauptsätze der allgemeinen Idealtheorie, Math. Annalen 90, S. 55–65, abgeleitet werden.

§ 3.

Reguläre Erweiterungen spezieller zerlegbarer Ringe.

Wir hatten in A. II § 7 einen speziellen Erweiterungsring $\bar{R}$ des speziellen Ringes R — wo $\bar{R}$ also auch wieder nur Nullteiler und Einheiten enthält und das System aller Nullteiler ein Primideal bildet — eine *reguläre* Erweiterung von R genannt, wenn $\bar{R}$ eine reguläre Modulbasis hinsichtlich R besitzt, d. h. wenn sich aus $\bar{R}$ eine (endliche oder unendliche) Menge von Elementen $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ herausgreifen läßt, derart, daß für jedes Element α aus $\bar{R}$ eine eindeutig bestimmte Darstellung $\alpha = \sum_{i=1}^{\nu} a_i \alpha_{\tau_i}$ durch endlich viele der α_τ mit Koeffizienten a aus R gilt, wobei insbesondere bei einem Nullteiler α ausschließlich Nullteiler aus R als Koeffizienten auftreten. Im Bereich der zerlegbaren Ringe gilt nun für reguläre Erweiterungen folgender grundlegender Doppelsatz:

Satz 5. *Jede reguläre Erweiterung $\bar{R}$ eines speziellen zerlegbaren Ringes R ist selbst wieder ein spezieller zerlegbarer Ring, und zwar ist jeder Primteiler aus R auch in $\bar{R}$ Primteiler. Ist umgekehrt $\bar{R}$ ein spezieller zerlegbarer Erweiterungsring des speziellen zerlegbaren Ringes R und ist ein Primteiler aus R auch Primteiler in $\bar{R}$, so ist $\bar{R}$ eine reguläre Erweiterung von R.*

Der erste Teil unserer Behauptung folgt aus der Definition der regulären Modulbasis einfach so: Es sei $\alpha = \sum_{i=1}^{\nu} a_i \alpha_{\tau_i}$ ein beliebiges Element aus $\bar{R}$, p sei ein beliebiger Primteiler aus R. Dann stellt jedes Element a_i das Produkt einer Potenz von p mit einer Einheit aus R dar, und wir können daher σ so bestimmen, daß ν Gleichungen $a_i = p^\sigma \cdot b_i$ $(i = 1, 2, \ldots, \nu)$ bestehen, wobei die b_i Elemente aus R und nicht sämtlich Nullteiler sind. Dann aber haben wir: $\alpha = p^\sigma \cdot \sum_{i=1}^{\nu} b_i \alpha_{\tau_i}$ und das Element $\sum_{i=1}^{\nu} b_i \alpha_{\tau_i}$ muß gemäß der Definition der regulären Erweiterung eine Einheit aus $\bar{R}$ sein. Es stellt also jedes Element aus $\bar{R}$ das Produkt einer Potenz von p mit einer Einheit dar, d. h. $\bar{R}$ ist ein spezieller zerlegbarer Ring mit dem Primteiler p.

Nehmen wir nun umgekehrt von vornherein an, daß $\bar{R}$ ein spezieller zerlegbarer Ring und p gleichzeitig in R und $\bar{R}$ Primteiler ist, so können wir zunächst folgendes bemerken: Sind die Elemente $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ aus $\bar{R}$ linear unabhängig, d. h. besteht zwischen endlich vielen von ihnen keine Relation $\sum_{i=1}^{\nu} a_i \alpha_{\tau_i} = 0$ mit nicht sämtlich verschwindenden Koeffizienten aus R, so ist ein Element von der Form $\sum_{i=1}^{\nu} a_i \alpha_{\tau_i}$ nur dann Nullteiler, wenn sämtliche a_i Nullteiler sind.

In der Tat, soll $\sum_{i=1}^{\nu} a_i \alpha_{\tau_i}$ ein Nullteiler sein, so haben wir unter unsern Vor. $\sum_{i=1}^{\nu} a_i \alpha_{\tau_i} = p^{\sigma}\cdot\beta;\ \sigma > 0$. Daraus folgt aber durch Multiplikation mit $p^{\varrho-\sigma} \neq 0$:[16]) $\sum_{i=1}^{\nu}(p^{\varrho-\sigma}\cdot a_i)\cdot\alpha_{\tau_i} = p^{\varrho}\cdot\beta = 0$. Wegen der linearen Unabhängigkeit der α_i folgt dann weiter $p^{\varrho-\sigma}\cdot a_i = 0$ $(i = 1, 2, \ldots, \nu)$, d. h. es müssen alle a_i durch p^{σ} teilbar, also gemäß unserer Behauptung Nullteiler sein.

Sind nun die Elemente $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ hinsichtlich R linear unabhängig, und ist α ein beliebiges Element aus $\overline{R}$, so ist entweder α von den α_τ linear unabhängig, oder es besteht eine Kongruenz $\alpha \equiv \sum_{i=1}^{\nu} a_i \alpha_{\tau_i}\ ((p))$. Soll nämlich die Gesamtheit der Elemente $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots, \alpha$ nicht linear unabhängig sein, so muß wegen der linearen Unabhängigkeit von $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ eine Gleichung $\sum_{i=1}^{\nu} a_i \alpha_{\tau_i} + a\cdot\alpha = 0$ mit von 0 verschiedenem a bestehen. Ist nun p^{σ} $(\sigma < \varrho)$ die höchste in $a_1, a_2, \ldots, a_\nu, a$ aufgehende Potenz von p, und setzen wir $a_i = p^{\sigma}\cdot b_i$, $a = p^{\sigma}\cdot b$, so folgt aus unserer Gleichung die Kongruenz $\sum_{i=1}^{\nu} b_i\cdot\alpha_{\tau_i} + b\cdot\alpha \equiv 0\,(p^{\varrho-\sigma})$.[17]) Wäre nun b ein Nullteiler, so müßte mindestens ein b_i eine Einheit sein, und es wäre wegen $\varrho - \sigma > 0$ das Element $\sum_{i=1}^{\nu} b_i \alpha_{\tau_i}$ durch p teilbar und folglich ein Nullteiler, was nach dem oben Festgestellten mit der linearen Unabhängigkeit von $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ im Widerspruch stände. Es muß mithin b eine Einheit sein, und wir haben $\alpha \equiv -\sum_{i=1}^{\nu}(b^{-1}\cdot b_i)\cdot\alpha_{\tau_i}\ ((p))$.

Mit Hilfe der beiden eben gemachten Bemerkungen sind wir nunmehr imstande, eine reguläre Modulbasis von $\overline{R}$ hinsichtlich R zu konstruieren. Es seien $\alpha_1, \alpha_2, \ldots, \alpha_\tau, \ldots$ sämtliche Elemente von $\overline{R}$ in einer beliebigen Wohlordnung[18]). Nur wollen wir der Einfachheit halber α_1 als Einheit voraussetzen. Wir zeigen dann durch transfinite Induktion: *Für jedes* τ *kann man aus den Elementen* α_i $(i \leqq \tau$ *bzw.* $i < \tau$, *falls* $\tau - 1$ *nicht existiert) eine Teilmenge* $\mathfrak{M}_\tau = \{\alpha_{\sigma_1} = \beta_1, \alpha_{\sigma_2} = \beta_2, \ldots\}$ *linear unabhängiger Elemente herausgreifen, derart, daß jedes* α_i *für* $i \leqq \tau$ *bzw.* $i < \tau$ *einer Kongruenz* $\alpha_i \equiv \sum_{i=1}^{\nu} a_i \beta_{\tau_i}\ ((p))$ *genügt. Dabei kann insbesondere* $\mathfrak{M}_\tau$ *stets*

[16]) Unter ϱ verstehen wir immer den charakteristischen Exponenten von R, so daß $p^{\varrho-1} \neq 0$; $p^{\varrho} = 0$ ist.

[17]) Man beachte, daß hier die Vor. benutzt wird, daß p einen Primteiler aus $\overline{R}$ bedeutet!

[18]) Falls $\tau - 1$ nicht existiert, entspricht der Ordnungszahl τ kein Element α_τ!

so gewählt werden, daß $\mathfrak{M}_\tau$ *die sämtlichen Mengen* $\mathfrak{M}_{\tau_1}$ $(\tau_1 < \tau)$ *enthält.* In der Tat, für $\tau = 1$ ist $\mathfrak{M}_1 = \{\alpha_1\}$ die gewünschte Menge, gibt es ferner für $\tau - 1$ eine Menge $\mathfrak{M}_{\tau-1}$, so können wir nach dem oben Festgestellten $\mathfrak{M}_\tau = \{\mathfrak{M}_{\tau-1}, \alpha_\tau\}$ bzw. $\mathfrak{M}_\tau = \mathfrak{M}_{\tau-1}$ setzen, je nachdem α_τ von den Elementen der Menge $\mathfrak{M}_{\tau-1}$ linear unabhängig ist oder nicht; und ist schließlich die Behauptung im Falle, daß $\tau - 1$ nicht existiert, für alle $\tau_1 < \tau$ richtig, so gilt sie auch für $\tau_1 = \tau$, denn man braucht für $\mathfrak{M}_\tau$ dann nur die Vereinigungsmenge aller $\mathfrak{M}_{\tau_1}$ $(\tau_1 < \tau)$ zu wählen[19]).

Verstehen wir nun unter $\mathfrak{M} = \{\beta_1, \beta_2, \ldots, \beta_\tau, \ldots\}$ die Vereinigungsmenge aller $\mathfrak{M}_\tau$, so sind die Elemente β_τ ersichtlich alle linear unabhängig, und es kann daher, wie oben festgestellt, $\beta = \sum_{i=1}^{\nu} a_i \beta_{\tau_i}$ nur dann Nullteiler sein, wenn alle a_i Nullteiler sind. *Daraus folgt, daß die Elemente* β_τ *eine reguläre Modulbasis von* $\bar{R}$ *hinsichtlich* R *bilden, sofern wir nur zeigen können, daß sich jedes Element aus* $\bar{R}$ *linear durch die* β_τ *darstellen läßt.* Das ergibt sich aber ohne Schwierigkeit aus dem in A. II so oft benutzten „Korrektionsgliederverfahren". In der Tat, ist α ein beliebiges Element aus $\bar{R}$, so haben wir angesichts der Konstruktion der Menge $\mathfrak{M}$ eine Gleichung $\alpha = \sum_{i=1}^{\nu_1} a_i^{(1)} \beta_{\tau_i^{(1)}} + p \cdot \alpha_1$. Indem wir jetzt α_1 modulo p durch die β_τ ausdrücken, kommen wir weiter zu einer Gleichung $\alpha = \sum_{i=1}^{\nu_2} a_i^{(2)} \beta_{\tau_i^{(2)}} + p^2 \cdot \alpha_2$ und nach ϱ Schritten finden wir: $\alpha = \sum_{i=1}^{\nu_\varrho} a_i^{(\varrho)} \beta_{\tau_i^{(\varrho)}}$. Die Elemente β_τ bilden also wirklich eine reguläre Modulbasis von $\bar{R}$ hinsichtlich R, es gilt nicht nur der erste Teil von Satz 5, sondern auch die Umkehrung.

Die für Satz 5 wesentliche Bedingung, daß R *und* $\bar{R}$ *den gleichen Primteiler* p *besitzen, ist mit der Forderung identisch, daß der charakteristische Exponent* ϱ' *von* $\bar{R}$ *gleich dem charakteristischen Exponenten* ϱ *von* R *ist.* In der Tat, ist p' ein Primteiler aus $\bar{R}$, p ein solcher aus R, so ist $p = r \cdot p'^{\lambda}$, wobei r eine Einheit bedeutet, und wegen $p^{\varrho-1} \neq 0$; $p^\varrho = 0$ ergibt sich daraus für ϱ die Ungleichung: $(\varrho - 1)\cdot\lambda < \varrho' \leqq \varrho\cdot\lambda$. Daraus folgt, daß ϱ dann und nur dann gleich ϱ' ist, wenn die Gleichung $\lambda = 1$ gilt, d. h. wenn p auch in $\bar{R}$ Primteiler ist.

Auf Grund der gemachten Bemerkung können wir Satz 5 kurz so aussprechen:

Die regulären Erweiterungen des speziellen zerlegbaren Ringes R *sind die einzigen speziellen zerlegbaren Erweiterungsringe mit dem gleichen charakteristischen Exponenten.*

---

[19]) Vgl. zu diesem Beweis den ganz analogen Beweis von Satz 21, A. II § 8!

Aus dem gewonnenen Ergebnis folgt insbesondere:

Ist R ein spezieller zerlegbarer Ring mit dem Primteiler p, so ist jede regulär algebraische und transzendente Erweiterung $\overline{R}$ ein spezieller zerlegbarer Ring mit dem gleichen Primteiler.

Denn die regulär algebraischen und transzendenten Erweiterungen besitzen ja, wie in A. II § 7 gezeigt wurde, stets eine reguläre Modulbasis hinsichtlich des Ausgangsringes.

Zum Schluß dieses Paragraphen wollen wir noch einen Satz beweisen, aus dem hervorgeht, daß die in A. I und A. II gewonnenen Resultate bei zerlegbaren Ringen in wesentlich einfacherer Weise abgeleitet werden können, als es im umfassenden Fall möglich war. Wir zeigen nämlich:

Satz 6. *Ist α ein Element aus einer beliebigen, speziellen Erweiterung des speziellen zerlegbaren Ringes R, so ist α hinsichtlich R regulär algebraisch, wenn es Nullstelle einer Primfunktion aus R_f ist, es ist hinsichtlich R regulär transzendent, wenn es Nullstelle keiner regulären Funktion aus R_f ist.*

Um den ersten Teil unseres Satzes zu beweisen, müssen wir zeigen, daß α als Nullstelle von $g(x)$ keiner Gleichung von niedrigerem Grade als $g(x)$ mit Koeffizienten aus R genügen kann. In der Tat, ist $a(x) = 0$ eine solche Gleichung, so können wir $a(x)$ auf die Gestalt $a(x) = p^{\mu} \cdot f(x)$ bringen, wobei $f(x)$ eine reguläre Funktion von niedrigerem Grade als $g(x)$ bedeutet. Dann aber müssen $f(x)$ und $g(x)$ teilerfremd sein, und aus $g(\alpha) = 0$, $a(\alpha) = 0$ folgt $p^{\mu} = 0$, d. h. $a(x)$ muß identisch verschwinden. — In ähnlicher Weise zeigen wir zum Beweise des zweiten Teiles unseres Satzes, daß α notwendig Nullstelle eines regulären Polynoms sein muß, falls es überhaupt einer nicht identisch bestehenden Gleichung mit Koeffizienten aus R genügt. Ist nämlich $a(\alpha) = 0$, so setzen wir wie oben $a(x) = p^{\mu} \cdot f(x)$, wobei $f(x)$ ein reguläres Polynom bedeutet. Sollte nun $p^{\mu} \neq 0$ sein, so ist $f(\alpha)$ ein Nullteiler des Erweiterungsringes, dem α angehört. Da aber dieser Erweiterungsring nach Vor. speziell ist, so haben wir für eine gewisse natürliche Zahl ϱ': $f^{\varrho'}(\alpha) = 0$, d. h. α ist Nullstelle des regulären Polynoms $f^{\varrho'}(x)$.

Satz 6 konnte im umfassenden Fall nur unter der einschränkenden Vor. bewiesen werden, daß α einer *regulären* Erweiterung von R angehört. Man könnte nun vielleicht auf Grund des hier gewonnenen weitergehenden Resultats annehmen, daß sich etwa die Theorie der unvollkommenen zerlegbaren Ringe wesentlich gründlicher ausbauen ließe, als es im umfassenden Fall möglich war. Daß dem nicht so ist, lehren die Gegenbeispiele, durch die wir in A. I und A. II zeigten, daß man bei

unvollkommenen Ringen nicht zu den gleichen Isomorphiekriterien gelangen kann wie bei den vollkommenen; denn diese Gegenbeispiele sind dem Gebiet der zerlegbaren Ringe entnommen.

§ 4.

Allgemeine einfache algebraische Erweiterungen zerlegbarer Ringe.

Wir stellen uns in diesem Paragraphen die Aufgabe, den allgemeinsten Fall zu untersuchen, in dem wir von dem speziellen zerlegbaren Ringe R durch Adjunktion eines einzigen Elementes α wiederum zu einem speziellen zerlegbaren Ringe $R(\alpha)$ gelangen. Ohne Interesse ist die Möglichkeit, daß α Nullstelle von überhaupt keiner Funktion aus $R_{\mathfrak{f}}$ ist. Denn dann gelangen wir offenbar zu einer regulär transzendenten Erweiterung von R, wenn wir die Gesamtheit aller rationalen Funktionen in α mit Koeffizienten aus R bilden, bei denen das Nennerpolynom mindestens einen regulären Koeffizienten besitzt.

Wir beschäftigen uns daher in Zukunft nur mit den Fällen, in denen α Nullstelle mindestens eines Polynoms aus $R_{\mathfrak{f}}$ ist — es soll $R(\alpha)$ dann als *allgemeine algebraische Erweiterung* von R bezeichnet werden. *Die Gesamtheit der Polynome in x, die für $x = \alpha$ verschwinden, bilden ein Ideal $\mathfrak{a}$ aus $R_{\mathfrak{f}}$, und dieses Ideal muß, wie aus dem Satz* 6 *des vorangehenden Paragraphen folgt, mindestens eine reguläre Funktion enthalten.* Ferner ist ohne weiteres klar, daß in $\mathfrak{a}$ kein von 0 verschiedenes Element aus R vorkommen kann, denn ein Element aus R soll ja in dem Erweiterungsring $\bar{R}$ nur dann verschwinden, wenn es bereits in R verschwindet. *Schließlich erkennt man leicht, daß das Ideal $\mathfrak{a}$ primär sein muß, falls $R(\alpha)$ ein spezieller zerlegbarer Ring sein soll.*

Wäre nämlich $\mathfrak{a}$ nicht primär, so wäre es das Produkt von zwei teilerfremden Idealen $\mathfrak{a}_1$ und $\mathfrak{a}_2$. Dann aber könnte man aus $\mathfrak{a}_1$ und $\mathfrak{a}_2$ zwei Funktionen $f_1(x)$ und $f_2(x)$ auswählen, die teilerfremd, und daher sicher nicht durch $\mathfrak{a}$ teilbar wären. Es gäbe mithin wegen $f_1(x) \cdot f_2(x) \equiv 0\ (\mathfrak{a})$ und der daraus folgenden Gleichung $f_1(\alpha) \cdot f_2(\alpha) = 0$ in $R(\alpha)$ zwei teilerfremde Nullteiler, nämlich $f_1(\alpha)$ und $f_2(\alpha)$, $R(\alpha)$ wäre also sicher nicht speziell. Als primäres Ideal muß $\mathfrak{a}$ nach den Entwicklungen von § 2 eine gekürzte Normalbasis

$$\mathfrak{a} = (h(x),\, p^{\mu_1} \cdot h_1(x),\, p^{\mu_1+\mu_2} \cdot h_2(x),\, \ldots,\, p^{\mu_1+\mu_2+\cdots+\mu_\sigma} \cdot h_\sigma(x))$$

besitzen, wobei p einen Primteiler aus R bedeutet, und die Funktionen $h(x) = h_0(x), h_1(x), \ldots, h_\sigma(x)$ sämtlich zum Primideal $\mathfrak{p} = (g(x), p)$ ge-

hörige Primärfunktionen sind[20]), mithin die Gestalt $h_i(x) = l_i(x)\cdot g_i(x)$ besitzen.

Wir behandeln zunächst den Fall, daß $\overline{R} = R(\alpha)$ denselben charakteristischen Exponenten wie R hat, und daß infolgedessen ein Primteiler p aus R auch in $\overline{R}$ Primteiler ist. Aus den Entwicklungen von § 3 kann man leicht schließen, daß in diesem Falle $\overline{R}$ eine regulär algebraische Erweiterung von R darstellen muß, man kann aber diese Tatsache auch ohne die oben angestellten allgemeinen Überlegungen einfach so einsehen:

Da das Element $g(\alpha)$ offenbar in $\overline{R}$ ein Nullteiler sein muß, und da p auch in $\overline{R}$ Primteilereigenschaft besitzen soll, so muß eine Gleichung $p^{\mu}\cdot f(\alpha) = g(\alpha)$ bestehen, wobei $f(\alpha)$ ein reguläres Element aus $\overline{R}$ bedeutet, und diese Gleichung ist mit der Kongruenz $p^{\mu}\cdot f(x) \equiv g(x)$ $(\mathfrak{a})$ identisch. $\mathfrak{a}$ enthält also die Primfunktion $g(x) - p^{\mu}\cdot f(x)$, und muß mit dem durch diese Primfunktion erzeugten Ideale $(g(x))$ identisch sein, weil andernfalls in $\mathfrak{a}$ entgegen unserer Voraussetzung ein von 0 verschiedenes Element aus R aufträte[21]).

Der Ring $\overline{R}$ ist demnach dem Restklassensystem nach einer Primfunktion aus R_f isomorph, d. h. er ist eine einfache reguläre algebraische Erweiterung von R, und wir können den Satz aussprechen:

Satz 7. *Ist α nicht hinsichtlich R transzendent und ist $R(\alpha)$ ein spezieller zerlegbarer Ring mit den gleichen charakteristischen Exponenten wie R, so ist $R(\alpha)$ eine regulär algebraische Erweiterung des Ringes R.*

In einfacher Weise läßt sich die Frage nach der allgemeinsten algebraischen Erweiterung noch in dem Falle beantworten, daß R ein Körper ist, also den charakteristischen Exponenten $\varrho = 1$ besitzt. Hier hat jedes primäre Ideal $\mathfrak{a}$ aus R_f die Gestalt $(g(x))^{\varrho'}$, wobei $g(x)$ eine Primfunktion bedeutet, und man sieht unmittelbar, daß der durch Adjunktion einer Nullstelle α von $\mathfrak{a}$ entstehende Ring $\overline{R} = R(\alpha)$ einen zerlegbaren Ring mit dem charakteristischen Exponenten ϱ' und dem Primteiler $g(\alpha)$ darstellt.

Satz 8. *Die allgemeinste einfache algebraische Erweiterung, die von einem Körper R zu einem speziellen zerlegbaren Ring mit dem charakteristischen Exponenten ϱ' führt, entsteht aus R durch Adjunktion einer Nullstelle der ϱ'-ten Potenz einer Primfunktion $g(x)$ aus R_f, d. h. sie ist zu dem Restklassensystem nach $(g(x)^{\varrho'})$ isomorph.*

[20]) Man beachte hier, daß die Möglichkeit, daß eine der Funktionen $h_i(x)$ eine Einheit ist, durch die Voraussetzung, $\mathfrak{a}$ solle kein Element aus R enthalten, ausgeschlossen ist!

[21]) Man beachte, daß $g(x)$ als Primfunktion zu jeder regulären Funktion niedrigeren Grades teilerfremd ist.

Wir betrachten jetzt den allgemeinen Fall, daß R einen von 0 verschiedenen Primteiler p enthält, der im Ringe $\bar{R}$ die Primteilereigenschaft verliert. Es sei dann p' ein Primteiler aus $\bar{R}$. Als Nullteiler muß p' die Gestalt $p' = g(\alpha)\cdot k(\alpha) + p\cdot k_1(\alpha)$ besitzen, wobei wir $k_1(\alpha)$ von niedrigerem Grade als $g(\alpha)$ annehmen dürfen. Wäre nämlich $p' = g(\alpha)\cdot k(\alpha) + r(\alpha)$, wobei $r(\alpha)$ eine reguläre Funktion von niedrigerem Grade als $g(x)$ bedeutet, so wären $h(x) = l(x)\cdot g(x) + q(x)$ und $g(x)\cdot k(x) + r(x)$ teilerfremd.

Zwischen p und p' muß eine Gleichung von der Form $p\cdot m(\alpha) = p'^{\lambda}\ (\lambda > 1)$ bestehen, wobei $m(\alpha)$ ein reguläres Element aus $\bar{R}$ bedeutet. Aus diesem Grunde müssen $m(x)$ und $g(x)$ teilerfremd sein und $m(x)$ besitzt daher die Form $m(x) = m_1(x)\cdot g(x) + r(x)$, wobei $r(x)$ eine reguläre Funktion bedeutet, deren Ordnung niedriger ist als der Grad von $g(x)$. Die zwischen p und p' bestehende Gleichung ist identisch mit einer Kongruenz von der Form

$$(1)\qquad p\cdot m_1(x)\cdot g(x) + p\cdot r(x) - F(x)\cdot g(x) - p^{\lambda}\cdot k_1(x)^{\lambda} \equiv 0(\mathfrak{a}),$$

dabei bezeichnet $F(x)\cdot g(x)$ den durch $g(x)$ teilbaren Bestandteil von $(k(x)\cdot g(x) + p\cdot k_1(x))^{\lambda}$. Ist nun $\mathfrak{a}$ durch das Ideal $(g(x), p^2)$ teilbar, so geht aus der Kongruenz (1) die Kongruenz

$$(2)\qquad p\cdot r(x) - p^{\lambda}\cdot k_1(x)^{\lambda} \equiv 0((g(x), p^2))$$

hervor, und daraus folgt weiter wegen $\lambda > 1$

$$(3)\qquad p\cdot r(x) \equiv 0\ ((g(x), p^2)).$$

Die Kongruenz (3) ist aber unmöglich, weil $r(x)$ eine reguläre Funktion von niedrigerem Grade als $g(x)$ ist. *Als notwendige Bedingung dafür, daß $\bar{R}$ einen speziellen zerlegbaren Ring darstellt, ergibt sich also, daß das Ideal $\mathfrak{a}$ und mithin nach einer oben gemachten Bemerkung alle durch $\mathfrak{a}$ teilbaren Funktionen niedrigsten Grades nicht durch das Ideal $(g(x), p^2)$, also auch nicht durch das Ideal $(g(x), p)^2$ teilbar sein dürfen.*

Es ist nun zu untersuchen, wie weit diese Bedingung auch hinreicht. Zunächst betrachten wir den Fall, daß $\mathfrak{a} = (h(x))$ ein Hauptideal ist. Jeder Nullteiler des Ringes $\bar{R}$ besitzt dann die Gestalt $k(\alpha)\cdot g(\alpha) + p^{\sigma}\cdot r(\alpha)$, wobei $r(\alpha)$ eine reguläre Funktion von niedrigerem Grade als $g(\alpha)$ bedeutet, und umgekehrt ist jedes Element dieser Form ein Nullteiler. Der Beweis dieser Tatsache folgt mühelos daraus, daß $h(x)$ eine zum Primideal $(g(x), p)$ gehörige Primärfunktion ist. Machen wir jetzt die Annahme, daß $h(x)$ nicht durch das Ideal $(g(x), p^2)$ teilbar ist, so muß $h(x)$ die Gestalt $h(x) = l(x)\cdot g(x) + q(x)$ besitzen, wobei $q(x)$ von niedrigerem Grade als $g(x)$ und nicht durch p^2 teilbar ist, d. h. wir

haben $q(x) = p \cdot r_1(x)$, wobei $r_1(x)$ eine reguläre Funktion bedeutet, die von niedrigerem Grade als $g(x)$, und mithin zu $g(x)$ und $h(x)$ teilerfremd ist. Da nun $h(x)$ als Primärfunktion die Gestalt $h(x) = e(x) \cdot g(x)^\mu + Q(x)$ besitzen muß, so ergibt sich aus der Gleichung $h(\alpha) = 0$ eine Gleichung von der Form $g(\alpha)^\mu = p \cdot e^{-1}(\alpha) \cdot (F(\alpha) \cdot g(\alpha) + r_1(\alpha))$, und es muß $F(\alpha) \cdot g(\alpha) + r_1(\alpha)$ eine Einheit aus $\bar{R}$ sein, weil, wie oben bemerkt, $h(x)$ und $r_1(x)$, also auch $h(x)$ und $F(x) \cdot g(x) + r_1(x)$ teilerfremd sind. Daher gilt in $\bar{R}$ eine Gleichung $p = g(\alpha)^\mu \cdot G(\alpha)$ und jeder Nullteiler aus $\bar{R}$ läßt sich als Produkt einer Einheit mit einer Potenz von $g(\alpha)$ darstellen, d. h. *$\bar{R}$ ist ein spezieller zerlegbarer Ring mit dem Primteiler $g(\alpha)$*. Es handelt sich um die Bestimmung des zu $\bar{R}$ gehörigen charakteristischen Exponenten ϱ'. Jedenfalls ist $g(\alpha)^{\mu \cdot \varrho} = 0$, d. h. es ist $\varrho' \leqq \mu \cdot \varrho$. Wäre aber $\varrho' < \mu \cdot \varrho$, so wäre wegen des Bestehens der Gleichung $p = g(\alpha)^\mu \cdot G(\alpha)$ auch $p^{\nu_1} \cdot g(\alpha)^{\nu_2} = 0$, wenn man

$$\varrho' = \nu_1 \cdot \mu + \mu_2 \quad (\nu_1 < \varrho, \nu_2 < \mu)$$

setzte. Das ist aber unmöglich, weil $g(x)^{\nu_2} \cdot p^{\nu_1}$ von niedrigerem Grade als $h(x)$ und $p^{\nu_1} \neq 0$ ist. *Wenn also $\bar{R}$ das Restklassensystem nach einer Primärfunktion darstellt, so ist sein charakteristischer Exponent durch $\mu \cdot \varrho$ gegeben.*

Wir wenden uns jetzt zur Behandlung des allgemeinsten Falles. Die bisher angestellten Überlegungen, durch die wir zeigten, daß $\bar{R}$ ein spezieller zerlegbarer Ring ist, falls $\mathfrak{a}$ nicht durch $(g(x), p)^2$ teilbar ist, lassen sich unmittelbar hierauf übertragen. Wir haben noch zu untersuchen, welchen Beschränkungen das Ideal $\mathfrak{a}$ dadurch unterworfen ist, daß wir fordern, es solle keine Elemente aus R enthalten. Zu diesem Zwecke betrachten wir eine gekürzte Normalbasis von $\mathfrak{a}$:

$$(4) \qquad \mathfrak{a} = (h(x), p^{\mu_1} \cdot h_1(x), p^{\mu_1+\mu_2} \cdot h_2(x) \ldots).$$

Dabei dürfen wir annehmen, daß in (4) allgemein $\mu_i > 0$ ist, und daß die $h_i(x)$ zur Primfunktion $g(x)$ gehörige Primärfunktionen von niedrigerem Grade als $h(x)$ bedeuten. Ist $h_1(x) = g(x)^\nu + q_1(x)$ $(\mu > \nu)$, so besteht infolge der speziellen Gestalt von $h(x)$ eine Kongruenz von der Form

$$(5) \qquad h_1(x) \equiv g(x)^\mu \cdot k_1(x) + g(x)^\nu \quad ((h(x))).$$

(Man beachte, daß nach den für den Fall $\mathfrak{a} = (h(x))$ angestellten Überlegungen eine Kongruenz $p \equiv g(x)^\mu \cdot G(x)$ $((h(x)))$ gelten muß.)

Da nun die Funktionen $h(x)$ und $g(x)^{\mu-\nu} \cdot k_1(x) + r_\varepsilon$ teilerfremd sind, so gibt es eine reguläre Funktion $k_2(x)$, die der Kongruenz $k_2(x) \cdot h_1(x) \equiv g(x)^\nu$ $((h(x)))$ Genüge leistet. Man kann also in der

Basis von $\mathfrak{a}$ die Funktion $p^{\mu_1} \cdot h_1(x)$ durch $p^{\mu_1} \cdot g(x)^\nu$ ersetzen. Ferner muß $\mu_1 = \varrho - 1$ sein; denn aus der Gleichung $p^{\mu_1} \cdot g(\alpha)^\nu = 0$ folgt $p^{\mu_1} \cdot g(\alpha)^\mu = 0$ und wegen des Bestehens der Gleichung $h(\alpha) = 0$ ergibt sich hieraus wegen der Gleichheit der Ideale (p) und $(g(\alpha)^\mu)$ die Gleichung $p^{\mu_1+1} = 0$. Es ist also gemäß unserer Behauptung entweder $\mu_1 = \varrho$, und dann ist $\mathfrak{a}$ ein Hauptideal, oder es ist $\mu_1 = \varrho - 1$, und $\mathfrak{a}$ besitzt die Basis $(h(x), p^{\varrho-1} \cdot g(x)^\nu)$ $(0 < \nu < \mu)$. Im letzteren Falle ist der charakteristische Exponent von $\bar{R}$ durch $\varrho' = \mu \cdot (\varrho - 1) + \nu$ gegeben. Das Gesamtresultat unserer Untersuchung über nicht reguläre algebraische Erweiterungen spezieller zerlegbarer Ringe fassen wir unter Beachtung der Bemerkung vom Schlusse von § 2 in dem folgenden Satze zusammen:

Satz 9. *Die allgemeinste einfache algebraische Erweiterung, die von dem speziellen zerlegbaren Ring R mit dem charakteristischen Exponenten $\varrho \geqq 2$ zu einem gleichfalls zerlegbaren speziellen Ringe $\bar{R}$ führt und die Eigenschaft besitzt, daß alle in R verschiedenen Elemente auch in $\bar{R}$ verschieden sind, besteht aus den Restklassen aller Polynome einer Variablen α mit Koeffizienten aus R nach einem nicht durch das Quadrat seines Primideals teilbaren Primärideal $\mathfrak{a}$ von der Form $(g(x)^\mu + q(x), p^{\varrho-1} g(x)^\nu)$, wobei $g(x)$ eine Primfunktion aus dem zu $\mathfrak{a}$ gehörigen Primideal bedeutet, und die Ungleichung $0 < \nu \leqq \mu$ besteht. Die niedrigste Potenz, für die der Primteiler $p' = g(\alpha)$ aus $\bar{R}$ verschwindet, ist $\varrho' = \mu(\varrho - 1) + \nu$.*[22])

Eine spezielle Folgerung aus dem eben formulierten Theorem ist nachstehender Satz, der in einem Spezialfalle bereits von Herrn Fraenkel bewiesen wurde[23]).

Satz 10. *Ist R ein spezieller zerlegbarer Ring mit dem Exponenten $\varrho = 2$, so ist jedes Restklassensystem nach einer irreduziblen Funktion aus R_f gleichfalls ein spezieller zerlegbarer Ring.*

In der Tat, auf jeden Fall muß eine irreduzible Funktion Primärfunktion sein. Unter unseren Voraussetzungen ist aber ferner diese Primärfunktion entweder eine Primfunktion, oder sie genügt der in Satz 8 geforderten Bedingung, weil man sonst von ihr eine Primfunktion als Faktor abspalten könnte. Wir bemerken noch:

Das Restklassensystem nach einer reduziblen Primärfunktion ist nie ein spezieller zerlegbarer Ring, wenn der Ausgangsring kein Körper ist.

Ist nämlich $h(x) = (g(x)^{\nu_1} + q_1(x)) \cdot (g(x)^{\nu_2} + q_2(x))$, so sind beide Faktoren durch $(g(x), p)$ und folglich $h(x)$ durch $(g(x), p)^2$ teilbar.

---

[22]) Besitzt das Ideal $\mathfrak{a}$ eine eingliedrige Basis, so ist $\nu = \mu$ zu setzen!

[23]) Über gewisse Teilbereiche und Erweiterungen von Ringen. Leipzig b. Teubner (1916), S. 57.

§ 5.

Fortgesetzte Betrachtung der allgemeinen algebraischen Erweiterungen.

Wir wollen nun die allgemeinen algebraischen Erweiterungen noch etwas eingehender untersuchen. Dabei müssen wir uns, wie durch ein Beispiel gezeigt werden soll, um zu befriedigenden Ergebnissen zu kommen, ähnlich wie in A. I § 10 auf vollkommene Ringe beschränken.

Satz 11. *Ist $\overline{R}$ eine allgemeine einfache algebraische Erweiterung des vollkommenen Ringes R, so können wir zwischen R und $\overline{R}$ einen Ring $\overline{\overline{R}}$ derart einschalten, daß $\overline{\overline{R}}$ eine regulär algebraische Erweiterung von R darstellt, während $\overline{R}$ aus $\overline{\overline{R}}$ einfach durch Adjunktion eines neuen Primteilers (also durch Adjunktion einer Nullstelle eines zum Primideal $(x, \mathfrak{p}^*)$ gehörigen Primideals) entsteht.*

Ist umgekehrt $\overline{\overline{R}}$ eine regulär algebraische Erweiterung von R und entsteht $\overline{R}$ aus $\overline{\overline{R}}$ durch Adjunktion eines neuen Primteilers, so kann $\overline{R}$ als einfache allgemeine algebraische Erweiterung von R aufgefaßt werden.

Es sei $\overline{R} = R(\alpha)$, wobei α eine Nullstelle des zum Primideal $(g(x), \mathfrak{p}^*)$ gehörigen Primärideals $\mathfrak{a}$ bedeuten möge. Dann ist $g(\alpha)$ ein Nullteiler aus $\overline{R}$, den wir mit $\overline{\mathfrak{p}}$ bezeichnen wollen, es ist mithin α Nullstelle der Funktion $g(x) - \overline{\mathfrak{p}}$ aus $\overline{R}_f$. Es bezeichne nun K den dem Ringe R, $\overline{K}$ den dem Ringe $\overline{R}$ im Sinne von A. I § 6 zugeordneten Körper[24]), $\overline{g}(x)$ bedeute das den modulo $\mathfrak{p}^*$ kongruenten Funktionen $g(x)$ und $g(x) - \overline{\mathfrak{p}}$ zugeordnete Polynom aus K_f. $\overline{g}(x)$ ist in K_f irreduzibel und muß wegen der Vollkommenheit von K im Körper $\overline{K}$, der eine algebraische Erweiterung von K darstellt, in lauter teilerfremde irreduzible Faktoren zerfallen. Daraus ergibt sich nach A. I § 7, daß die Funktionen $g(x)$ und $g(x) - \overline{\mathfrak{p}}$ in $\overline{K}_f$ in teilerfremde Primfunktionen zerfallen müssen, und zwar müssen bei geeigneter Zuordnung entsprechende Primfunktionen der beiden Zerlegungen modulo $\mathfrak{p}_f^*$ kongruent sein. Daraus ergibt sich, daß $g(x)$ in $\overline{R}$ eine mit α modulo $\mathfrak{p}^*$ kongruente Nullstelle α' besitzen muß. α' ist nach Satz 6 hinsichtlich R regulär algebraisch, der Ring $\overline{\overline{R}} = R(\alpha')$ stellt eine reguläre Erweiterung von R dar, und wegen der Kongruenz $\alpha \equiv \alpha' (\mathfrak{p}^*)$ ist den Ringen $\overline{\overline{R}}$ und $\overline{R}$ derselbe Körper $\overline{K}$ zugeordnet, d. h. $\overline{\overline{R}}$ enthält aus jeder Klasse modulo $\mathfrak{p}^*$ kongruenter Elemente von $\overline{R}$ mindestens eines. Adjungieren wir nun zu $\overline{\overline{R}}$ einen beliebigen Primteiler $\overline{\mathfrak{p}}$ aus $\overline{R}$, so kommen wir zu einem Bereiche $\overline{R}'$, dem erstens der Körper $\overline{K}$ zugeordnet ist, und der zweitens aus jeder Klasse äquivalenter Primteiler von $\overline{R}$ mindestens einen Vertreter enthält. Aus diesen

24) Also den Körper, der aus R bzw. $\overline{R}$ dadurch entsteht, daß man alle Nullteiler dem Nullelement gleichsetzt.

beiden Tatsachen ergibt sich aber, nach dem in A. I und A. II häufig benutzten „Korrektionsgliederverfahren“ [25]), die Identität von $\overline{R}$ und $\overline{R}'$, und hiermit ist der erste Teil von Satz 11 bewiesen.

Zum Beweis des Schlusses nehmen wir an, es sei α ein Element, durch dessen Adjunktion $\overline{\overline{R}}$ aus R entsteht, während $\overline{p}$ einen Primteiler aus $\overline{R}$ bedeuten möge. Dann ist $\overline{R} = R(\alpha + \overline{p})$. Sicher enthält nämlich $R(\alpha + \overline{p})$ aus jeder Klasse modulo $\mathfrak{p}^*$ kongruenter Elemente von $\overline{R}$ mindestens einen Vertreter. Wir haben also, wie genau wie oben aus dem Korrektionsgliederverfahren folgt, zum Beweise der Gleichung $\overline{R} = R(\alpha + \overline{p})$ nur zu zeigen, daß $R(\alpha + \overline{p})$ einen Primteiler aus $\overline{R}$ enthalten muß. Das ist aber der Fall, wir haben nämlich nach der Taylorentwicklung:

$$g(\alpha + \overline{p}) = g(\alpha) + \overline{p} \cdot g'(\alpha) + \overline{p}^2 \cdot h(\alpha) = \overline{p} \cdot (g'(\alpha) + \overline{p} \cdot h(\alpha)),$$

wobei $g'(x)$ die formal gebildete Ableitung von $g(x)$ bedeutet. Wegen der Vollkommenheit des Ringes R ist nun $g'(x)$ zu $g(x)$ teilerfremd, und es ist daher $g'(\alpha)$, also $g'(\alpha) + \overline{p} \cdot h(\alpha)$ ein reguläres Element aus $\overline{R}$, es stellt mithin $\overline{p} \cdot (g'(\alpha) + \overline{p} \cdot h(\alpha))$ einen in $R(\alpha + \overline{p})$ auftretenden Primteiler dar, und daraus folgt, wie oben bemerkt, die Gleichung $\overline{R} = R(\alpha + \overline{p})$. Der Beweis von Satz 14 ist hiermit abgeschlossen.

Wir wollen noch durch ein Gegenbeispiel zeigen, daß mindestens der erste Teil des Satzes für unvollkommene Ringe seine Gültigkeit verliert. Es sei $R = K_2(t)$ der Körper aller rationalen Funktionen von t mit Koeffizienten modulo 2, und es sei weiter $\overline{R} = R(\alpha)$, wobei α eine Nullstelle des Ideals $((x^2 - t)^2)$ bedeutet. Dann enthält $\overline{R}$ kein hinsichtlich R regulär algebraisches, von den Elementen von R selbst verschiedenes Element. Andernfalls müßte nämlich insbesondere in $\overline{R}$ ein Element $\beta = \sqrt{t}$ vorkommen, das der Gleichung $\beta^2 - t = 0$ genügte [26]), und dieses Element müßte mit α modulo $\mathfrak{p}^*$ kongruent sein, weil ja α modulo $\mathfrak{p}^*$ ebenfalls Nullstelle des Polynoms $x^2 - t$ ist. Bedeutet aber q einen Nullteiler, so haben wir wegen $2 \cdot r_e = 0$; $q^2 = 0$ stets $(\alpha + q)^2 = \alpha^2$, und wegen $\alpha^2 - t \neq 0$ kann daher ein Element β von der gewünschten Art in $\overline{R}$ nicht existieren. Bei unvollkommenen Ringen ist es also mitunter unmöglich, zwischen R und $\overline{R}$ einen hinsichtlich R regulär algebraischen Ring $\overline{\overline{R}}$ so einzuschalten, daß $\overline{R}$ aus $\overline{\overline{R}}$ durch Adjunktion eines Primteilers entsteht, Satz 11 verliert seine Gültigkeit.

Wir wenden uns wieder zu vollkommenen Ringen. Der Ring $\overline{R}$ soll als eine allgemeine endliche algebraische Erweiterung des Ringes R be-

---

[25]) Vgl. die Anmerkung bei § 6, wo nochmals von der hier benutzten Schlußweise Gebrauch gemacht wird.

[26]) Dies folgt einfach daraus, daß der dem Ringe $\overline{R}$ zugeordnete Körper $\overline{K}$ aus $R = K$ durch Adjunktion von $\sqrt{t}$ entsteht.

zeichnet werden, wenn $\bar{R}$ aus R durch sukzessive Ausführung von endlich vielen allgemeinen einfachen algebraischen Erweiterungen entsteht. Dann gilt

Satz 12. *Jede allgemeine endliche algebraische Erweiterung eines vollkommenen zerlegbaren Ringes ist einfach.*

Es sei R der Ausgangs-, $\bar{R}$ der Erweiterungsring. Dann genügt es nach dem letzten Teil von Satz 11 zum Beweise von Satz 12, wenn wir zeigen, daß sich zwischen R und $\bar{R}$ ein Erweiterungsring $\bar{\bar{R}}$ so einschalten läßt, daß $\bar{\bar{R}}$ eine einfache regulär algebraische Erweiterung von R darstellt, während $\bar{R}$ aus $\bar{\bar{R}}$ durch Adjunktion eines Primteilers erzeugt werden kann. Wir führen den Beweis dieser Tatsache der Einfachheit halber für den Fall, daß $\bar{R}$ aus R durch zweimalige allgemeine algebraische Erweiterung entsteht. Die Verallgemeinerung auf n-malige Erweiterung ist dann trivial. Es sei also $R^{(1)} = R(\alpha)$; $\bar{R} = R^{(1)}(\beta)$, wobei α ein hinsichtlich R, β ein hinsichtlich $R^{(1)}$ im allgemeinen Sinne algebraisches Element bedeutet. Dann können wir nach Satz 11 zwei Zwischenringe $\bar{\bar{R}}^{(1)}$ und $\bar{\bar{R}}^{(2)}$ so bestimmen, daß $\bar{\bar{R}}^{(1)}$ hinsichtlich R, $\bar{\bar{R}}^{(2)} = R^{(1)}(\beta')$ hinsichtlich $R^{(1)}$ regulär algebraisch ist, und daß $R^{(1)}$ aus $\bar{\bar{R}}^{(1)}$, $\bar{R}$ aus $\bar{\bar{R}}^{(2)}$ jeweils durch Adjunktion eines Primteilers entsteht. Ist nun $g(x)$ diejenige Primfunktion aus $R_{\mathfrak{f}}^{(1)}$, deren Nullstelle β' ist, so kann man eine zu $g(x)$ modulo $\mathfrak{p}_{\mathfrak{f}}^{*}$ kongruente Primfunktion $g'(x)$ aus $\bar{\bar{R}}_{\mathfrak{f}}^{(1)}$ finden, da ja $\bar{\bar{R}}^{(1)}$ aus jeder Klasse modulo $\mathfrak{p}^{*}$ kongruenter Elemente von $R^{(1)}$ mindestens eines enthält. Wegen der Vollkommenheit von $R^{(1)}$ und $\bar{\bar{R}}^{(1)}$ folgt nun nach bekannten Schlüssen, daß es in $\bar{\bar{R}}^{(2)}$ eine mit β' modulo $\mathfrak{p}^{*}$ kongruente Nullstelle β'' von $g'(x)$ gibt und daß β'' hinsichtlich $\bar{\bar{R}}^{(1)}$ regulär algebraisch ist. Der Ring $\bar{\bar{R}} = \bar{\bar{R}}^{(1)}(\beta'')$ ist mithin eine endliche regulär algebraische, also wegen der Vollkommenheit von R eine einfache reguläre Erweiterung von R. Ferner ist dem Ringe $\bar{\bar{R}}$, wie leicht (etwa durch Gradabzählung) einzusehen, derselbe Körper zugeordnet wie dem Ringe $\bar{\bar{R}}^{(2)}$ und folglich wie dem Ringe $\bar{R}$. $\bar{R}$ kann also aus $\bar{\bar{R}}$ durch Adjunktion eines Primteilers gewonnen werden. Hiermit ist, in Anbetracht von Satz 11, Satz 12 für eine zweifache Erweiterung bewiesen, und das Ergebnis läßt sich, wie sofort zu sehen, unmittelbar auf n-fache Erweiterungen übertragen.

Mit dem zuletzt gewonnenen Resultat wollen wir die Betrachtung der allgemeinen endlichen Erweiterungen abschließen. Aus den Sätzen 11 und 12 folgt, daß man, um tiefer in die Natur dieser Erweiterungen einzudringen, sich jedenfalls auf das Studium von solchen Erweiterungsringen beschränken darf, die aus dem Ausgangsring durch Adjunktion eines Primteilers entstehen; denn die regulär algebraischen Erweiterungen eines vollkommenen Ringes wurden ja bereits in A. I § 10 erschöpfend behandelt.

Man könnte auch noch allgemeine unendliche Erweiterungen in Betracht ziehen. Diese dürften in verschiedener Hinsicht von großem Interesse sein, sie führen aber jedenfalls zu Bereichen, die keine zerlegbaren Ringe im Sinne von § 1 mehr sind. Das Studium der allgemeinen unendlichen Erweiterungen fällt daher aus dem der vorliegenden Arbeit gesteckten Rahmen heraus.

§ 6.
Die endlichen zerlegbaren Ringe.

Um sämtliche Typen von endlichen zerlegbaren Ringen zu erhalten, genügt es natürlich, die Typen der speziellen zerlegbaren Ringe aufzustellen, da man ja aus diesen nach der im zweiten Teil entwickelten Methode die allgemeinen ableiten kann.

Was nun die speziellen zerlegbaren Ringe angeht, so ist zunächst zu bemerken, *daß jeder (endliche und nicht endliche) „Grundring“*[27]) *einen zerlegbaren Ring darstellt.* Denn ein Grundring entsteht ja aus dem in evidenter Weise zerlegbaren Ringe der Restklassen nach einer Primzahlpotenz durch regulär algebraische und transzendente Erweiterungen, ist also nach § 3 selbst zerlegbar, und zwar gibt es unter den Vielfachen des Einheitselements einen Primteiler des Grundringes.

Es handelt sich nun darum, zu untersuchen, wie ein beliebig gegebener spezieller zerlegbarer Ring aus seinem Grundring entsteht. Er ist jedenfalls dann mit seinem Grundring identisch, wenn im Grundring ein Primteiler vorkommt; im anderen Falle kann er aus dem Grundring durch Adjunktion eines Primteilers erzeugt werden. Ist nämlich R_1 ein Teilbereich von R, der den Grundring sowie einen Primteiler von R enthält, so enthält R_1 aus jeder Klasse modulo $\mathfrak{p}^*$ kongruenter Elemente sowie aus jeder Klasse äquivalenter Nullteiler mindestens einen Vertreter und muß daher nach einer im ersten und zweiten Teile verschiedentlich angewandten Schlußweise mit R identisch sein[28]).

Es sei jetzt R ein von seinem Grundring $R^{(1)}$ verschiedener, zerlegbarer Ring, p sei einer seiner Primteiler. Dann muß p hinsichtlich $R^{(1)}$ algebraisch sein und die Adjunktion von p zu $R^{(1)}$ stellt daher eine der in den beiden vorangehenden Paragraphen besprochenen, allgemeinen algebraischen Erweiterungen dar. Die in § 4 und § 5 hergeleiteten Sätze können uns daher sofort zur Aufstellung eines Theorems über die endlichen zerlegbaren Ringe dienen. Wir haben nur die Tatsache zu beachten, daß für die in Satz 8 und Satz 9 auftretende Primfunktion $g(x)$ im vorliegenden Falle x gewählt werden darf, da ja p der Gleichung $p^{\varrho} = 0$ genügt.

[27]) Zum Begriffe des Grundringes vgl. A. II § 4 u. 6.

[28]) Vgl. A. I § 10 S. 120, sowie A. II § 7.

Satz 13. *Es sei R ein endlicher, spezieller zerlegbarer Ring, p einer seiner Primteiler, ϱ sein charakteristischer Exponent, während p_0 und ϱ_0 für den zugehörigen Grundring $R^{(1)}$ dieselbe Bedeutung haben. Dann sind folgende Fälle möglich:*

a) $\varrho = \varrho_0$. *In diesem Falle ist $R^{(1)}$ mit R identisch.*

b) $\varrho = \mu \cdot (\varrho_0 - 1) + \nu$ $(\mu > 1;\ 0 < \nu \leqq \mu)$. *In diesem Falle entsteht R aus $R^{(1)}$ durch Adjunktion von p und dasjenige Ideal aus dem zu $R^{(1)}$ gehörigen Funktionenring, dessen Nullstelle p ist, besitzt eine Basis von der Form $(x^\mu + q(x),\ p_0^{\varrho_0 - 1} \cdot x^\nu)$, wobei $q(x)$ für $p_0 \neq 0$ nicht durch p_0^2 teilbar ist.*

Ferner ergibt sich aus den Resultaten von § 5 folgender Satz, der in der Theorie der Ideale eines algebraischen Zahlkörpers eine bedeutende Rolle spielt[29]):

Satz 14. *Jeder endliche spezielle zerlegbare Ring kann aus dem (aus den Vielfachen des Einheitselements bestehenden) Primring $R^{(0)}$ durch einfache algebraische Erweiterung gewonnen werden.*

Satz 14 ist eine unmittelbare Folge des zweiten Teils von Satz 12. Denn der Grundring ist ja eine endliche regulär algebraische Erweiterung des Primrings, während R selbst, wie eben festgestellt, aus seinem Grundring durch Adjunktion eines Primteilers entsteht.

Über die Anzahl der Elemente eines endlichen speziellen zerlegbaren Ringes gibt der folgende Satz Aufschluß:

Satz 15. *Ist ϱ der charakteristische Exponent des endlichen speziellen zerlegbaren Ringes R, und gibt es genau π^σ Klassen modulo $\mathfrak{p}^*$ inkongruenter Elemente aus R (π Primzahl)[30]), so enthält R genau $\pi^{\sigma \cdot \varrho}$ Elemente, unter denen sich $\pi^{\sigma \cdot \varrho} \cdot (1 - \pi^{-\sigma})$ reguläre befinden.*

Es bedeute nämlich p einen Primteiler aus R, S einen Bereich, der aus jeder Klasse modulo $\mathfrak{p}^*$ inkongruenter Elemente genau eines enthält. Dann erhält man alle Elemente aus R und jedes nur einmal, wenn man alle Polynome von der Form $\sum\limits_{i=0}^{\varrho-1} a_i p^i$ mit Koeffizienten aus S betrachtet. Ebenso lassen sich alle Nullteiler aus R eindeutig in der Form $\sum\limits_{i=1}^{\varrho-1} a_i p^i$ darstellen.

---

[29]) Vgl. z. B. Hilbert, Die Theorie der algebraischen Zahlkörper. Jahresbericht d. Deutschen Mathematikervereinigung 4 (1897), S. 193.

[30]) Die Anzahl der mod $\mathfrak{p}^*$ inkongruenten Elemente aus R ist sicher eine Primzahlpotenz, sie ist nämlich gleich der Anzahl der Elemente des dem Ringe R zugeordneten endlichen Körpers K.

Jedes reguläre Element aus dem endlichen speziellen zerlegbaren Ringe R genügt der Gleichung $x^{\pi^{\sigma\cdot\varrho}(1-\pi^{-\sigma})} - r_\varepsilon = 0$. Doch hat dieses Analogon zum kleinen Fermatschen Satz keine tiefere Bedeutung, da die Funktion $x^{\pi^{\sigma\cdot\varrho}(1-\pi^{-\sigma})} - r_\varepsilon$ nur dann Primfunktion ist, wenn R einen Körper darstellt.

Bekanntlich sind zwei endliche Körper isomorph, wenn sie dieselbe Elementezahl besitzen. *Entsprechendes gilt von zwei Grundringen, die in der Anzahl ihrer Elemente und im charakteristischen Exponenten übereinstimmen.* Man könnte vermuten, daß bei den endlichen speziellen zerlegbaren Ringen die Verhältnisse genau so liegen wie bei den Grundringen. Daß dem aber nicht so ist, zeigt das folgende Beispiel.

Als Grundring wählen wir das Restklassensystem modulo 8, und erzeugen aus diesem Grundring durch Adjunktion von $\sqrt{2}$ bzw. $\sqrt{6}$ zwei spezielle zerlegbare Ringe R_1 und R_2 mit den Primteilern $\sqrt{2}$ bzw. $\sqrt{6}$. Wäre nun R_1 zu R_2 isomorph, so müßte R_1 neben $\sqrt{2}$ auch $\sqrt{6}$ enthalten. Das kann aber nicht sein, denn sonst käme in R_1 — da $\sqrt{2}$ und $\sqrt{6}$ Primteiler sind, sich also nur um eine Einheit unterscheiden können — auch $\sqrt{6} : \sqrt{2} = \sqrt{3}$ vor, und das ist, wie leicht einzusehen, unmöglich, weil der Grundring $\sqrt{3}$ nicht enthält, da ja 3 quadratischer Nichtrest nach 8 ist. *R_1 und R_2 sind daher nicht isomorph, trotzdem sie in der Elementezahl und im charakteristischen Exponenten übereinstimmen.*

Als Beispiel für einen endlichen zerlegbaren Ring möge das System S der Restklassen nach einem beliebigen Ideale $\mathfrak{A}$ aus einem algebraischen Zahlkörper betrachtet werden. Zunächst haben wir uns zu überzeugen, daß S tatsächlich einen zerlegbaren Ring darstellt. Das ist der Fall, denn:

1. *S ist ein kommutativer Ring mit Einheitselement.*

2. *S enthält nur endlich viele Elemente, weil es nach $\mathfrak{A}$ nur endlich viele Restklassen gibt. Jedes Element aus S ist daher nach den Ergebnissen von Teil* II[31]) *entweder Einheit oder Nullteiler.*

3. *In S ist jedes Ideal ein Hauptideal, denn nach einem bekannten Satz wird in einem algebraischen Zahlkörper modulo einem festen Ideale jedes andere zum Hauptideal.*

Es sei jetzt $\mathfrak{A} = \prod_{i=1}^{\sigma} \mathfrak{P}_i^{\varrho_i}$, wobei die $\mathfrak{P}_i$ Primideale bedeuten, dann ist eine beliebige Restklasse aus S ein Nullteiler, wenn einer, und folglich jeder ihrer Repräsentanten durch eines der Ideale $\mathfrak{P}_i$ teilbar ist, im andern Fall ist die Restklasse eine Einheit. Zwei Restklassen sind dann und nur dann teilerfremd, wenn nicht die Repräsentanten von beiden gleichzeitig durch eines der Ideale $\mathfrak{P}_i$ teilbar sind. *Aus diesen Tatsachen ergibt sich*

---

31) Vgl. A. II § 4.

leicht, daß die speziellen zerlegbaren Ringe, auf die man S zurückführen kann, durch die Restklassensysteme nach $\mathfrak{P}_1^{\varrho_1}, \mathfrak{P}_2^{\varrho_2}, \ldots, \mathfrak{P}_\sigma^{\varrho_\sigma}$ *dargestellt werden.*

Wir betrachten also das Restklassensystem R nach der ϱ-ten Potenz eines Primideals $\mathfrak{P}$. Offenbar ist ϱ der charakteristische Exponent von R und ein Element p aus R ist dann und nur dann Primteiler, wenn seine Repräsentanten durch $\mathfrak{P}$, aber nicht durch $\mathfrak{P}^2$ teilbar sind. Um nun zu ermitteln, welcher der in Satz 13 charakterisierten Typen durch R dargestellt wird, beachten wir, daß es eine natürliche Ringzahl π gibt, die durch $\mathfrak{P}$ teilbar ist, und zwar möge $\pi \equiv 0(\mathfrak{P}^\mu)$, $\pi \not\equiv 0(\mathfrak{P}^{\mu+1})$ sein. Die durch π repräsentierte Restklasse bezeichnen wir mit p_0. Sie stellt sicher einen Primteiler in dem zu R gehörigen Grundring $R^{(1)}$ dar. Je nach dem Verhältnis der Zahlen μ und ϱ ergeben sich uns folgende Fälle (mit ϱ_0 wird der charakteristische Exponent von $R^{(1)}$ bezeichnet):

a) $\mu = 1$. *In diesem Falle ist R mit seinem Grundring* $R^{(1)}$ *identisch, denn* p_0 *stellt einen Primteiler sowohl in* $R^{(1)}$ *als auch in R dar.*

b) $\mu > 1$. (*Dieser Fall tritt bekanntlich dann und nur dann ein, wenn* $\mathfrak{P}$ *in der Diskriminante des algebraischen Zahlkörpers, dem es entnommen ist, aufgeht.*) *Setzt man* $\varrho = \mu(\varrho_0 - 1) + \nu$ $(0 < \nu \leqq \mu)$, *so ist durch diese Gleichung* ϱ_0 *eindeutig bestimmt, und es stellt den charakteristischen Exponenten von* $R^{(1)}$ *dar*, wie man sofort erkennt, wenn man sich überlegt, welche Potenz von π unter unsern Voraussetzungen durch $\mathfrak{P}^\varrho$ teilbar wird, also verschwindet. — *R entsteht aus* $R^{(1)}$ *durch Adjunktion von p und* $\mathfrak{P}$ *ist Nullstelle eines Ideals* $(x^\mu + q(x),\ p_0^{\varrho_0 - 1} x^\nu)$ *aus* $R_f^{(1)}$.

Da es stets Primideale $\mathfrak{P}$ gibt, die in der Diskriminante ihres Zahlkörpers, also in ihrer Primzahl π in einer höheren als der ersten Potenz aufgehen, so lassen sich, wie man aus der eben durchgeführten Diskussion erkennt, Beispiele für alle die in Satz 11 aufgezählten Fälle bereits unter den Restklassensystemen nach Potenzen von Primidealen aus algebraischen Zahlkörpern finden, der allgemeinste endliche zerlegbare Ring ist also nicht wesentlich komplizierter gebaut als das Restklassensystem nach einem geeignet gewählten algebraischen Ideal. Mit dem eben Festgestellten ist aber noch nicht nachgewiesen, daß sich jeder endliche zerlegbare Ring durch ein derartiges Restklassensystem verkörpern läßt. Um über die Möglichkeit oder Unmöglichkeit einer solchen Darstellung entscheiden zu können, müßte man zunächst folgendes zahlentheoretische Problem lösen:

Es seien $\pi_1, \pi_2, \ldots, \pi_n$ *beliebige untereinander verschiedene Primzahlen,* σ_{ik}, μ_{ik} $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m_i)$ *seien beliebige natürliche Zahlen. Gibt es dann stets einen algebraischen Zahlkörper, in dem für die Primzahlen* π_i *eine Zerlegung* $\pi_i = \prod_{k=1}^{m_i} \mathfrak{P}_k^{\mu_{ik}} \cdot \mathfrak{Q}_i$ *gilt, wobei die*

*$\mathfrak{P}_{ik}$ zueinander und zu $\mathfrak{Q}_i$ teilerfremde Primideale bedeuten, und wobei insbesondere das Restklassensystem nach dem Primideal $\mathfrak{P}_{ik}$ genau $\pi_i^{\sigma_{ik}}$ verschiedene Elemente enthält?**) Läßt sich die aufgeworfene Frage bejahen, so ist damit folgendes gezeigt:

Es seien π_i $(i = 1, 2, \ldots, n)$ beliebige, verschiedene natürliche Primzahlen, σ_{ik}, $\varrho_{ik}^{(0)}$, ϱ_{ik} $(i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, m_i;\ \varrho_{ik} \geqq \varrho_{ik}^{(0)})$ seien beliebige natürliche Zahlen. Dann läßt sich stets ein algebraisches Ideal $\mathfrak{A}$ bestimmen, so daß das Restklassensystem nach $\mathfrak{A}$ einen endlichen zerlegbaren Ring S bildet, der sich als Summe von $\sum_{i=1}^{n} m_i$ speziellen Ringen $R^{(ik)}$ $(i = 1, 2, \ldots, r;\ k = 1, 2, \ldots, m_i)$ darstellen läßt, wobei $R^{(ik)}$ den charakteristischen Exponenten ϱ_{ik} und einen Grundring mit $\pi_i^{\sigma_{ik}}$ verschiedenen Elementen und mit dem Grundringexponenten $\varrho_{ik}^{(0)}$ besitzt.

Bestimmt man nämlich μ_{ik} so, daß die Gleichung

$$\varrho_{ik} = \mu_{ik} \cdot (\varrho_{ik}^{(0)} - 1) + \nu_{ik} \qquad (0 < \nu_{ik} \leqq \mu_{ik})$$

besteht, und sucht man dann einen algebraischen Zahlkörper auf, in dem die Primzahlen π_i in der oben angegebenen Weise in Idealfaktoren zerfallen, so stellt nach dem früher gewonnenen Ergebnis das Restklassensystem nach $\prod_{i=1}^{n} \prod_{k=1}^{m_i} \mathfrak{P}_{ik}^{\varrho_{ik}}$ den gewünschten Ring S dar.

Wäre nun ein endlicher spezieller Ring durch Grundring und charakteristischen Exponenten eindeutig bestimmt, so wäre mit der Lösung des oben formulierten zahlentheoretischen Problems die aufgeworfene Frage erledigt. Da aber, wie gezeigt, zu gegebenem Grundring und charakteristischem Exponenten mitunter mehrere wesentlich verschiedene Ringe existieren, so sind zur vollen Erledigung des vorgelegten Problems noch andere zahlentheoretische Untersuchungen als die oben angegebenen nötig. Hier genügt es, auf den ganzen Fragenkomplex nur hinzuweisen. Denn mit Hilfe von Satz 13 beherrschen wir den Aufbau eines endlichen zerlegbaren Ringes vollständig, wir brauchen daher zur Typisierung dieser Bereiche die Darstellung durch das Restklassensystem nach einem algebraischen Ideal gar nicht.

---

*) Vgl. die während des Druckes erschienene Abhandlung: „*Zur Theorie der Eisensteinschen Gleichungen*" von Herrn Oynstein Ore (Math. Zeitschrift **20**, S. 267–280), in deren § 4 die hier aufgeworfene Frage für $n = 1$ in bejahendem Sinne beantwortet ist.

§ 7.

Typisierung der allgemeinen vollkommenen zerlegbaren Ringe.

Zum Abschluß der vorliegenden Arbeit sollen noch einige Bemerkungen über die Typisierung von beliebigen vollkommenen zerlegbaren Ringen gemacht werden. Da bei einem vollkommenen Ringe die Existenz des Grundrings nach den Resultaten von A II feststeht, so können wir sagen:

Satz 13a. *Jeder vollkommene zerlegbare Ring kann aus seinem Grundring in derselben Weise wie ein endlicher zerlegbarer Ring durch allgemeine algebraische Erweiterung erzeugt werden.*

Mit Hilfe von Satz 13a können wir den Aufbau der vollkommenen zerlegbaren Ringe in genau der gleichen Weise übersehen wie den der endlichen Ringe.

Es sollen nun noch einige weitergehende Untersuchungen angestellt werden, die sich vorwiegend auf algebraisch abgeschlossene Ringe beziehen. Dabei wollen wir folgende Ausdrucks- und Schreibweise gebrauchen:

Es sei R der gegebene Ring. Dann bezeichnet $R^{(1)}$ seinen Grundring, ϱ den charakteristischen Exponenten von R, $\varrho^{(1)}$ denjenigen von $R^{(1)}$, den „Grundringexponenten". Ist $R^{(1)}$ kein Körper, gibt es also eine Primzahl π, so daß das Element $\pi \cdot r_e$ einen von Null verschiedenen Primteiler aus $R^{(1)}$ darstellt, so soll π als „kritische Primzahl" von R bezeichnet werden. Für das Element $\pi \cdot r_e$ wollen wir alsdann stets die Schreibweise p_1 benutzen, während $p, p', \ldots$ Primteiler aus R selbst bedeuten sollen. Als „kritische Invariante" μ des Ringes R bezeichnen wir schließlich die durch die Gleichung $\varrho = (\varrho_1 - 1)\cdot \mu + \nu$ $(0 < \nu \leqq \mu)$ eindeutig bestimmte positive Zahl μ.

Wir gehen jetzt zur Behandlung der Isomorphiefrage bei vollkommenen zerlegbaren Ringen über. Natürlich ist, wie das im vorangehenden Paragraphen gegebene Beispiel zeigt, ein vollkommener Ring durch Grundring und charakteristischen Exponenten nicht eindeutig bestimmt. Doch sind zwei wichtige Fälle hervorzuheben, bei denen die Sache anders liegt.

Satz 16. *Ein vollkommener zerlegbarer Ring, dessen Grundring ein Körper ist, ist durch Grundring und charakteristischen Exponenten bis auf Isomorphie eindeutig bestimmt.*

Der eben formulierte Satz ist beinahe trivial, da nach Satz 13a ein vollkommener Ring R, dessen Grundring $R^{(1)}$ ein Körper ist, aus $R^{(1)}$ durch Adjunktion einer Nullstelle des Ideals (x^{ϱ}) entsteht. Zu einem interessanteren

Falle gelangen wir, wenn wir algebraisch abgeschlossene Ringe betrachten. Es ist natürlich in Anbetracht von Satz 16 nur der Fall zu untersuchen, in dem der Grundring kein Körper ist.

Satz 17. *Es sei R ein algebraisch abgeschlossener zerlegbarer Ring. Dann ist R in folgenden Fällen durch Grundring und charakteristischen Exponenten eindeutig bestimmt:*

α) Wenn $R^{(1)}$ ein Körper ist.

β) Wenn die kritische Invariante μ zur kritischen Primzahl π teilerfremd ist.

γ) Wenn $\varrho = \mu + 1$; $\varrho_1 = 2$ ist.

In allen übrigen Fällen kann man stets zwei nicht isomorphe Ringe mit gleichem Grundring und gleichem charakteristischen Exponenten angeben.

α) ist bereits durch Satz 16 erledigt. In den Fällen β) und γ) führt uns der Nachweis zum Ziel, daß wir R stets aus $R^{(1)}$ durch Adjunktion einer Nullstelle p des Ideals $(x^\mu - p_1,\ x^\nu \cdot p_1^{\varrho_1 - 1})$ erhalten können, denn das angeschriebene Ideal ist ja nur vom Grundring und von der kritischen Invariante, also wegen des Zusammenhangs zwischen μ und ϱ nur von $R^{(1)}$ und von ϱ abhängig. Um das Element p in der gewünschten Weise zu bestimmen, betrachte man einen beliebigen Primteiler p' aus R. p' ist Nullstelle eines Ideals $(x^\mu - p_1 \cdot f(x),\ x^\nu \cdot p_1^{\varrho_1 - 1})$ aus $R_f^{(1)}$, wobei $f(p')$ ein reguläres Element aus R bedeutet. Ist nun μ zur kritischen Primzahl π teilerfremd, so zerfällt in dem dem Ringe R zugeordneten, algebraisch abgeschlossenen Körper K die Funktion $x^\mu - \overline{f(p')}$ in lauter teilerfremde Linearfaktoren, und es gilt daher nach A I § 6 das gleiche von dem Polynom $x^\mu - f(p')$ aus R_f. Wir können infolgedessen ein reguläres Element a aus R so wählen, daß es der Gleichung $x^\mu - f(p') = 0$ genügt. Setzt man $p' = p \cdot a$, so ist $p = p' \cdot a^{-1}$ ebenfalls ein Primteiler, und zwar ein solcher, der die Gleichung $x^\mu - p_1 = 0$ befriedigt. Das Ideal, dessen Nullstelle p ist, besitzt daher die gewünschte Basis $(x^\mu - p_1,\ x^\nu \cdot p_1^{\varrho_1 - 1})$. Etwas anders muß man im Falle γ) schließen. Hier verschwindet das Produkt von p_1 mit einem beliebigen Nullteiler aus R, und das Ideal, dessen Nullstelle p' ist, besitzt daher eine Basis $(x^\mu - a \cdot p_1,\ p_1 \cdot x)$, wobei a ein modulo $\mathfrak{p}^*$ beliebig wählbares Element aus R bedeutet. Wegen der algebraischen Abgeschlossenheit von R darf man alsdann a so gewählt annehmen, daß die Gleichung $x^\mu - a = 0$ in R eine Nullstelle a_1 besitzt. Setzt man $p' = p \cdot a_1$, so ist p eine der gesuchten Nullstellen des Ideals $(x^\mu - p_1,\ x \cdot p_1)$. Die Fälle β) und γ) sind hiermit erledigt.

Es bleibt nur noch zu zeigen, daß für $\varrho \neq \mu + 1$; $\varrho_1 \geqq 2$ stets zu

einem gegebenen Grundring mit der kritischen Primzahl π zwei nicht isomorphe Ringe mit derselben kritischen, zu π nicht teilerfremden Invariante μ aufgebaut werden können. Wir setzen $\mu = \pi^{\sigma} \cdot \mu'$; $(\mu', \pi) = 1$ und betrachten die Ideale

$$\mathfrak{a} = (x^{\mu} - p_1,\ p_1^{\varrho_1 - 1} \cdot x^{\nu}) \quad \text{und} \quad \mathfrak{a}' = (x^{\mu} - p_1 \cdot (x + r_{\varepsilon}),\ p_1^{\varrho_1 - 1} \cdot x^{\nu}).$$

Es sei p eine Nullstelle von $\mathfrak{a}$, p' eine solche von $\mathfrak{a}'$;

$$R = R^{(1)}(p), \qquad R' = R^{(1)}(p').$$

Dann können R und R' nicht isomorph sein.

Andernfalls müßte nämlich R ebenso wie R' eine Nullstelle des Ideals $\mathfrak{a}'$ enthalten, die wir der Einfachheit halber gleichfalls mit p' bezeichnen wollen. Da p und p' Primteiler sind, so besteht eine Gleichung $p' = a \cdot p$. Wir haben also $a^{\mu} \cdot p^{\mu} - p_1 \cdot (p' + r_{\varepsilon}) = (a^{\mu} - (p' + r_{\varepsilon})) \cdot p_1 = 0$. Setzen wir mithin $a^{\mu} = r_{\varepsilon} + p' + q$, so muß q der Gleichung $q \cdot p_1 = 0$ genügen, und daraus ergibt sich wegen $\varrho_1 \geqq 2$; $\varrho > \mu + 1$, daß q mindestens durch p'^2 teilbar sein muß, wir haben also $a^{\mu} = r_{\varepsilon} + p' \cdot r$, wobei r eine Einheit bedeutet. Wir betrachten nun die Differenz $s = r_{\varepsilon} - a^{\mu}$. Es ist

$$s^{\pi^{\sigma}} = r_{\varepsilon} + (-1)^{\pi^{\sigma}} \cdot a^{\mu} + p_1 \cdot b = r_{\varepsilon} \cdot (1 + (-1)^{\pi^{\sigma}}) + (-1)^{\pi^{\sigma}} p' \cdot r + p_1 \cdot b.$$

Da nun $r_{\varepsilon} \cdot (1 + (-1)^{\pi^{\sigma}}) = \begin{cases} p_1 & (\pi = 2) \\ 0 & (\pi \neq 2) \end{cases}$ ebenso wie $p_1 \cdot b$ mindestens durch p'^2 teilbar ist, so ist $s^{\pi^{\sigma}} = (-1)^{\pi^{\sigma}} p' (r + p' \cdot c) = p' \cdot r_1$ das Produkt eines Primteilers mit einer Einheit, also selbst Primteiler. Enthielte mithin R gleichzeitig die Elemente p und p', so müßte es in R ein Element s geben, das π^{σ}-te Wurzel eines Primteilers wäre[32]). Da die Existenz eines solchen Elementes der Natur des Primteilers widerspricht, so kann p' nicht in R vorkommen, R und R' sind also nicht isomorph. Mit dem so gewonnenen Ergebnis ist der Beweis unseres Satzes abgeschlossen.

Aus Satz 17 ergibt sich, daß man zwei isomorphe Ringe erhält, wenn man die bei Besprechung der endlichen Ringe aufgebauten nicht isomorphen Bereiche $R_0(\sqrt{2})$ und $R_0(\sqrt{6})$ zu algebraisch abgeschlossenen Ringen ergänzt. Es ist also, wenn zwei nicht isomorphe Ringe mit gleichem Grundring und gleicher kritischer Invariante vorgelegt sind, wohl zu unterscheiden,

[32]) Man vergleiche das hier gegebene Beispiel mit dem am Schlusse von A. I angegebenen, durch das bewiesen wurde, daß bei unvollkommenen Ringen zwei regulär algebraische Erweiterungen desselben Ausgangsringes mit gleichem zugeordneten Körper nicht stets äquivalent sind. Die hier und dort benutzten Schlüsse laufen ganz parallel. Die aus der kritischen Primzahl entspringenden Schwierigkeiten sind in beiden Fällen im ganzen dieselben. Der einzige wesentliche Unterschied ist der, daß hier die Primteiler eine analoge Rolle spielen, wie in A. I die transzendenten Elemente.

ob der Grund für die Nichtisomorphie an der algebraischen Natur des Grundrings liegt, oder ob wir es mit einer Nichtisomorphie zu tun haben, die mit der kritischen Primzahl zusammenhängt, und durch regulär algebraische Erweiterung nicht behoben werden kann.

Im Anschluß an Satz 17 kann man sich noch folgende Frage vorlegen. Es seien R und R' zwei algebraisch abgeschlossene, nicht isomorphe Ringe, mit gleichem Grundring und gleichem charakterischen Exponenten. Wann kann man dann die Nichtisomorphie durch allgemeine algebraische Erweiterung beheben, d. h. wann kann man einen Ring $\bar{R}$ finden, der sowohl einen zu R als auch einen zu R' isomorphen Teilbereich enthält, und der sich als allgemeine algebraische Erweiterung von R (bzw. von R') auffassen läßt?

Die hier aufgeworfene Frage hängt, wie leicht zu sehen[33]), mit folgendem Problem aufs engste zusammen:

Es sei $h(x)$ eine Primärfunktion μ-ten Grades aus R_f, von der wir wegen der algebraischen Abgeschlossenheit von R ohne wesentliche Beschränkung voraussetzen dürfen, daß sie zum Primideal (x, p) gehört. Wann enthält dann R eine Nullstelle von $h(x)$ und wann kann man eine allgemeine algebraische Erweiterung $\bar{R}$ von R derart bestimmen, daß $h(x)$, wenn auch nicht in R, so doch in $\bar{R}$ eine Nullstelle besitzt?

Die eben gestellte Frage bedeutet in gewissem Sinne eine Verallgemeinerung des Problems, das uns zur Einführung der allgemeinen algebraischen Erweiterungen hinleitete. Dort betrachteten wir eine Primärfunktion $h(x)$, die nicht durch das Quadrat des zugehörigen Primideals teilbar war, und stellten fest, daß man einfach durch Adjunktion einer Nullstelle von $h(x)$ zum gewünschten Erweiterungsring kommt (falls nicht $h(x)$ bereits im Ausgangsring eine Nullstelle besitzt). Im allgemeinen Falle liegt die Sache verwickelter. Da würde die Adjunktion einer Nullstelle von $h(x)$ selbst im allgemeinen zu einem Erweiterungsring führen, der

[33]) Man vergleiche das beim Beweise von Satz 17 angeführte Beispiel von zwei nicht isomorphen Ringen R und R'. Dort wurde gezeigt: In einem Ring $\bar{R}$, der gleichzeitig einen zu R und einen zu R' isomorphen Teilbereich enthält, muß die π^{σ}-te Wurzel eines gewissen, zu einem Primteiler aus R äquivalenten Elementes vorkommen. Das heißt aber nichts anderes, als daß eine gewisse, zum Primideal (x, p) gehörige Primärfunktion π^{σ}-ten Grades aus R_f in R eine Nullstelle besitzen muß. Dabei ist in unserem besonderen Falle die in Betracht kommende Primärfunktion $h(x)$ nicht durch das Quadrat des Primideals (x, p) teilbar. Man kann hier also $\bar{R}$ unmittelbar durch Adjunktion einer Nullstelle von $h(x)$ aus R erzeugen. Der so aufgebaute Ring $\bar{R}$ enthält dann, wie leicht zu sehen, wirklich einen zu R' isomorphen Teilbereich, in diesem besonderen Falle kann man also die Nichtisomorphie von R und R' sicher durch algebraische Erweiterung heben. Nicht so einfach gestaltet sich, wie oben betont, die Sache im allgemeinen Fall.

nicht mehr zerlegbar wäre. Gleichwohl erscheint es durchaus nicht ausgeschlossen, daß man durch Adjunktion einer Nullstelle einer anderen, geeignet gewählten Primärfunktion einen Erweiterungsring der gewünschten Art aufbauen könnte. Doch dürfte die allgemeine Beantwortung der aufgeworfenen Frage nicht einfach sein. Wir müssen uns daher hier mit dem Hinweis begnügen, daß die Einführung der allgemeinen algebraischen Erweiterungen auf eine Reihe von interessanten, einer ausführlichen Behandlung wohl würdigen Problemen hinleitet.

Für die vorliegende Arbeit reichen die mit Satz 13a und Satz 17 gewonnenen Ergebnisse vollständig aus. Satz 13a leistet für die vollkommenen zerlegbaren Ringe dasselbe wie Satz 13 für die endlichen: er gibt uns einen befriedigenden Einblick in diese Bereiche. Satz 17 hingegen charakterisiert in erschöpfender Weise diejenigen Fälle, in denen sich ein vollkommener Ring durch Grundring und charakteristischen Exponenten eindeutig bestimmen läßt. Am interessantesten ist dabei die Erkenntnis von der großen Bedeutung, die die kritische Primzahl für den Aufbau der zerlegbaren Ringe besitzt. Hier zeigt sich am schärfsten der Zusammenhang mit der Körpertheorie (vgl. die Bemerkung am Schlusse der Einleitung), der die in der vorliegenden Arbeit durchgeführte algebraische Behandlung der zerlegbaren Ringe rechtfertigt.

(Eingegangen am 31. 12. 1923.)

Aufstellung einer Transformationstheorie für eine neue Klasse aufeinander abwickelbarer Flächen.

Von

Hans Jonas in Berlin-Steglitz.

Inhaltsübersicht.

Einleitung.

Die im folgenden entwickelte Transformationstheorie der Biegungsflächen vom Typus

$$\text{(I)} \qquad ds^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2),$$

als dessen einfachster Vertreter die Translationsfläche

$$z = \frac{3}{2}(x^{2/3} - y^{2/3})$$

anzusprechen ist, stützt sich auf eine von Tzitzéica [1]) gefundene Flächentransformation, die im Bereiche derjenigen Flächen gilt, für die das Krümmungsmaß proportional mit der 4. Potenz des Abstandes der Tangentialebene vom Koordinatenanfang ist. Tzitzéicas Transformation der durch die Beziehung $\frac{K}{W^4} = \text{konst.}$ charakterisierten Flächen gehört zu den *asymptotischen* Flächentransformationen, bei denen die gegebene und die transformierte Fläche die beiden Brennflächenmäntel einer W-Kongruenz [2]) bilden. Man gelangt zu ihr, wie ich in einer unlängst in den Annali di Matematica erschienenen Abhandlung [3]) gezeigt habe, durch Spezialisierung einer allgemeineren asymptotischen Transformation, die zwischen den Flächen derjenigen Klasse vermittelt, die bei Beziehung auf die Parameter der Asymptotenlinien durch das Bestehen der Relation $\left\{ {11 \atop 2} \right\} = \left\{ {22 \atop 1} \right\}$ ausgezeichnet ist. Die Anwendung des allgemeinen Bianchischen Kompositionstheorems für die Moutardschen Transformationen gestattete die Aufstellung eines Vertauschbarkeitssatzes, demzufolge sich die Tzitzéicaschen Transformationen, die ich unter Hervorhebung einer charakteristischen Konstanten mit T_n bezeichnet habe, zu geschlossenen viergliedrigen Zyklen zusammensetzen lassen. Als von ganz besonderer Bedeutung für die gegenwärtige Untersuchung erwies sich ein am Schluß der genannten Abhandlung gewonnenes Ergebnis: Es lassen sich je drei zu dem gleichen Werte von n gehörige Transformationen T_n in der Weise aneinanderreihen, daß die erste und die letzte der vier durch sie verbundenen Flächen ähnlich und in bezug auf den Koordinatenanfang ähnlich gelegen sind.

Verbindet man nun mit einer Fläche der Tzitzéicaschen Klasse eine zweite, die ihr durch Polarreziprozität bezüglich einer (reellen oder imagi-

[1]) Tzitzéica, Sur une nouvelle classe de surfaces. C. R. de l'Acad. des Sc. **150** (1910), S. 955 u. 1227.

[2]) Auf den Brennflächenmänteln einer W-Kongruenz entsprechen sich bekanntlich die Asymptotenlinien und damit gleichzeitig die konjugierten Systeme.

[3]) Jonas, Sopra una classe di trasformazioni asintotiche etc. Annali di Mat. (3) **30** (1921), S. 223.

nären) Einheitskugel um den Koordinatenanfang zugeordnet ist, so führt eine einfache Überlegung, die von der bekannten Weingartenschen Behandlung des Biegungsproblems [4]) Gebrauch macht, zu der bemerkenswerten Tatsache, daß sich die Flächen der durch (I) definierten isometrischen Klasse mittels der Tzitzéicaschen Flächen konstruieren lassen. Wie weit Tzitzéica selber seine Untersuchungen in dieser Richtung ausgedehnt hat, geht aus seinen Veröffentlichungen nicht hervor. Den einzigen Anhaltspunkt bieten die folgenden Zeilen, mit denen er die erste seiner beiden, 1908 und 1909 in den Rendiconti del Circolo Matematico di Palermo erschienenen Abhandlungen [5]) über die „neue Flächenklasse" einleitet: „L'étude de la déformation des surfaces tétraédrales

$$\text{(II)} \qquad A x^{2/3} + B y^{2/3} + C z^{2/3} = 1$$

conduit à la recherche des surfaces pour lesquelles la courbure totale est proportionnelle à la 4$^{\text{me}}$ puissance de la distance d'un point fixe au plan tangent." Zu bemerken ist hierzu aber einmal, daß sich der typische Ausdruck (I) für das Quadrat des Linienelements in dem erwähnten Zusammenhange *nicht* findet und ebensowenig in zwei zeitlich voraufgehenden, von Tzitzéica [6]) und von Egorof [7]) herrührenden Noten in den Comptes Rendus, die sich auf die Eigentümlichkeit der Flächen der tetraedralen Klasse (II) beziehen, isometrische Deformationen innerhalb dieser selben Flächenklasse zuzulassen [8]). Überdies ist anzunehmen, daß Tzitzéica, als er die angeführten Worte schrieb, noch nicht im Besitze der Transformation T_n gewesen ist; diese ist nämlich in den beiden Abhandlungen noch nicht enthalten, ist vielmehr erst Gegenstand der unter [1]) zitierten Noten vom Jahre 1910.

---

[4]) Siehe z. B. Bianchi, Lezioni di geom. diff. 2 (1903), Cap. XIX.

[5]) Tzitzéica, Sur une nouvelle classe de surfaces. Palermo Rend. 25 (1908), S. 180; 28 (1909), S. 210.

[6]) Tzitzéica, Sur la déformation de certaines surfaces liées aux surfaces du second degré. C. R. de l'Ac. des Sc. 128 (1899), S. 1276.

[7]) Égorov, Une classe nouvelle de surfaces algébriques. C. R. de l'Ac. des Sc. 132 (1901), S. 302.

[8]) Die beiden Noten beschränken sich im wesentlichen auf den Nachweis, daß innerhalb der tetraedralen Klasse (II) stetige Verbiegungen mit Erhaltung eines konjugierten Systems möglich sind. Betreffs der Reduktion des Linienelements auf die Form

$$ds^2 = \frac{9}{4}\varkappa^2 (u\,du \pm 2\,du\,dv + v\,dv^2)$$

($\varkappa$ ist eine durch die Koeffizienten A, B, C bestimmte Konstante) und der besonderen Stellung dieser stetigen Deformationen innerhalb einer *zwei*-parametrigen Biegungsuntergruppe sei verwiesen auf: Jonas, Über eine mit den Flächen 2. Grades zusammenhängende zweifach-unendliche Schar aufeinander abwickelbarer Flächen. Berl. Math. Ges. Ber. 22 (1923), S. 49.

Hinsichtlich unserer neuen, im Mittelpunkt der vorliegenden Arbeit stehenden Transformation Θ_n, vermöge deren man aus einer Biegungsfläche vom Typus (I) beliebig viele weitere solche Flächen gewinnt, sei hier zunächst festgestellt, daß ihre geometrischen Eigenschaften ihr eine durchaus selbständige Stellung neben der Tzitzéicaschen Transformation T_n sichern. Zwischen den Transformationen Θ_n und T_n besteht ein ähnliches Verhältnis wie zwischen der Transformation der auf die Paraboloide abwickelbaren Flächen und der Transformation der Flächen von konstanter Krümmung. Es lassen sich übrigens, wenn man das Biegungsproblem für das Linienelement (I) und das durch die klassischen Arbeiten von Thybaut, Calapso und Bianchi durchaus noch nicht erschöpfte Problem der Verbiegung der Paraboloide gemeinsam ins Auge faßt, sehr weitgehende Analogien aufdecken, und zwar, ohne daß irgendwelche Anzeichen auf einen inneren Zusammenhang zwischen den beiden Differentialgleichungen

$$\theta_{\alpha\beta} = e^{\theta} - e^{-2\theta} \quad \text{und} \quad \theta_{\alpha\beta} = e^{\theta} - e^{-\theta}$$

schließen ließen, deren Transformation den analytischen Kern der beiden Theorien ausmacht. Erwähnt sei, daß gleichzeitig und in eigentümlicher Wechselwirkung mit der vorliegenden Arbeit eine andere, ihr in geometrischer Hinsicht eng verwandte entstanden ist, in der ich die Transformationen der Biegungsflächen des gleichseitigen hyperbolischen Paraboloids behandelt habe[9]).

Die nachfolgenden Entwicklungen, deren Gliederung aus der voraufgeschickten Inhaltsangabe zu ersehen ist, sind auch dort, wo es sich um Formeln der Transformation T_n handelt, unabhängig und ohne Bezugnahme auf die eingangs erwähnte Abhandlung in den Annali di Matematica durchgeführt. Es schien an sich schon geboten, an Stelle des dort benutzten Umwegs über eine Transformation allgemeineren Charakters hier einen neuen, direkten Weg einzuschlagen. Überdies fiel der Umstand ins Gewicht, daß unsere Transformation Θ_n nicht schlechthin gleichwertig mit zwei simultanen Transformationen T_n und $T_{\frac{1}{n}}$ der beiden zur Tzitzéicaschen Klasse gehörigen Hilfsflächen ist; die Koordinaten der transformierten Hilfsflächen müssen nämlich erst durch Multiplikation mit gewissen Konstanten *normiert* werden, damit die gewünschte Beziehung zu der transformierten Biegungsfläche hergestellt wird. Infolgedessen ließen sich Änderungen in den Bezeichnungen nicht umgehen.

Die Transformation Θ_n der Flächen vom Linienelement (I) ist, wiewohl auch bei ihr Korrespondenz der Asymptotenlinien zwischen der gegebenen und der transformierten Fläche besteht, keine asymptotische Trans-

[9]) Jonas, Ricerche sulle trasformazioni delle superficie applicabili sul paraboloide iperbolico equilatero. Annali di Mat. (im Druck).

formation, sondern von komplizierterer Natur. Ein tieferes Eindringen in den geometrischen Zusammenhang wird durch eine Tatsache ermöglicht, die in eigenartiger Weise den Vertauschbarkeitssatz für die Transformation Θ_n ergänzt. Es ist das die Existenz *drei*gliedriger Transformationszyklen, die zu dem oben erwähnten Ergebnis am Schluß meiner früheren Arbeit in engster Beziehung steht. Auf Grund dieser Eigenschaft gelingt es, *die Transformation* Θ_n *in zwei aufeinanderfolgende asymptotische Transformationen aufzulösen*, d. h. zwischen der gegebenen und der transformierten Biegungsfläche eine dritte Fläche einzuschalten, deren Punkte mit den entsprechenden der beiden Biegungsflächen die Brennpunkte der Strahlen zweier W-Kongruenzen bilden. Dieser, soweit es sich um die geometrische Anschauung handelt, jedenfalls höchst einfache Sachverhalt erschien mir, zumal da er sich durchaus nicht mühelos feststellen ließ, eher überraschend als etwa im Hinblick auf bekannte Transformationstheorien selbstverständlich. Die Darstellung folgt hier genau dem Gang der ursprünglichen Untersuchung.

Das eben erwähnte Ergebnis, das unsere Transformationstheorie zu einem gewissen Abschluß bringt, darf insofern ein ganz besonderes Interesse beanspruchen, als es eine Erweiterung des von Bianchi verwendeten Prinzips der durch W-Kongruenzen vermittelten Transformation von Biegungsflächen bestimmter Typen darstellt. Die außerordentlichen Erfolge, die Bianchi selber im Bereiche der auf die Flächen 2. Grades abwickelbaren Flächen[10]) erzielte, sind der Anlaß zu, wie es scheint, allzu hoch gespannten Hoffnungen hinsichtlich der Leistungsfähigkeit dieses Prinzips gewesen. Jedenfalls hat das nächstliegende Problem, die Bianchische Methode auf die Biegungsflächen aller bzw. allgemeinerer Regelflächen auszudehnen, bis jetzt allen Versuchen beharrlich widerstanden, ja, es mehren sich neuerdings die Ergebnisse negativer Art, die darauf schließen lassen, daß die angestrebte Verallgemeinerung auch nicht annähernd in dem erwarteten Umfange gelingen wird. Um so bedeutungsvoller erscheint die Tatsache, daß man imstande ist, *ohne den Kreis der Bianchischen Ideen zu verlassen, lediglich durch Paarung der asymptotischen Transformationen ein Seitenstück zur Transformationstheorie der auf die Flächen* 2. *Grades abwickelbaren Flächen aufzubauen.*

In betreff der speziellen Biegungsflächen vom Typus (I), die der tetraedralen Flächenklasse (II) angehören, ist zu bemerken, daß sie in analytischer Hinsicht einen Ausartungsfall bilden und dementsprechend bei der Aufstellung der Transformation Θ_n auszuschließen waren. Auf den Nachweis, daß sich die gepaarten asymptotischen Transformationen auch

---

[10]) Siehe z. B. Bianchi, Lezioni di geometria diff. 3 (1909).

zwischen diesen Flächen verwirklichen lassen, wurde mit Rücksicht auf den Umfang der Arbeit verzichtet.

Hingewiesen sei schließlich auf die Beziehungen der durch $\frac{K}{W^4} = \text{konst.}$ charakterisierten Tzitzéicaschen Flächen zur *affinen Geometrie*. Eine engere Anpassung der ganzen Untersuchung an diesen jüngsten Zweig der Differentialgeometrie, wie sie vielleicht zeitgemäß gewesen wäre, erschien mir bei der metrischen Natur des im Mittelpunkte stehenden Biegungsproblems als unvorteilhaft.

§ 1.

Eigenschaften der Biegungsflächen vom Typus $ds^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2)$; Einführung der beiden zugehörigen Tzitzéicaschen Flächen, einer dritten Hilfsfläche und einer Schar von Flächen $(x^{(\nu)}, \ldots)$.

1. Es sei (x, y, z) eine Fläche, für die das Quadrat des Linienelements die Form

$$ds^2 = \sum dx^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2) \tag{1}$$ [11])

besitzt, also eine Biegungsfläche[12]) der Translationsfläche

$$x = (u-1)^{3/2}, \quad y = (v-1)^{3/2}, \quad z = \frac{3}{2}(u - v),$$

deren erzeugende Kurven in zueinander senkrechten Ebenen liegen und deren Koordinatengleichung

$$z = \frac{3}{2}(x^{2/3} - y^{2/3}) \tag{2}$$

lautet.

Wir bilden für die quadratische Differentialform (1) die Christoffelschen Symbole

$$\left\{\begin{aligned} &\begin{Bmatrix}11\\1\end{Bmatrix} = \frac{v}{2(uv-1)}, \quad \begin{Bmatrix}11\\2\end{Bmatrix} = \frac{1}{2(uv-1)}, \quad \begin{Bmatrix}12\\1\end{Bmatrix} = \begin{Bmatrix}12\\2\end{Bmatrix} = 0, \\ &\begin{Bmatrix}22\\1\end{Bmatrix} = \frac{1}{2(uv-1)}, \quad \begin{Bmatrix}22\\2\end{Bmatrix} = \frac{u}{2(uv-1)} \end{aligned}\right. \tag{3}$$

[11]) Die Hinzufügung des Faktors $\frac{9}{4}$ erweist sich erst später als zweckmäßig. Sie gestattet die Unterdrückung eines sonst auftretenden Faktors $\frac{2}{3}$ in einer Reihe von Formeln.

[12]) Die Ausdrücke *Biegungsfläche*, *abwickelbar* und *Isometrie* sind durchweg im allgemeinsten Sinne zu verstehen. Sie beziehen sich lediglich auf die formale Übereinstimmung im Bau des Linienelements, ohne daß gemeinsame Wertbereiche der Variablen ins Auge gefaßt werden.

und finden für das Krümmungsmaß den Ausdruck:

(4) $$K = -\frac{1}{9(uv-1)^2}.$$

Mit X, Y, Z seien die Normalenkosinus, mit L, M, N die Fundamentalgrößen 2. Ordnung bezeichnet, so daß also

$$-\sum dx\,dX = L\,du^2 + 2M\,du\,dv + N\,dv^2$$

ist. Mit Berücksichtigung von (3) erhält man:

(5) $$x_{uu} = \frac{vx_u + x_v}{2(uv-1)} + LX, \quad x_{uv} = MX, \quad x_{vv} = \frac{x_u + ux_v}{2(uv-1)} + NX$$ [13])

und analoge Beziehungen für y und z.[14]) Die Gaußsche Relation und die Codazzischen Gleichungen haben die folgende Gestalt:

(6) $$\begin{cases} LN - M^2 = -\dfrac{9}{16(uv-1)}, \\ L_v - M_u + \dfrac{v}{2(uv-1)}M + \dfrac{N}{2(uv-1)} = 0, \\ N_u - M_v + \dfrac{u}{2(uv-1)}M + \dfrac{L}{2(uv-1)} = 0. \end{cases}$$

Ein Lösungssystem L, M, N von (6) definiert intrinsek eine Biegungsfläche (x, y, z) vom Typus (1).

2. Wir konstruieren nun zwei Hilfsflächen (ξ, η, ζ) und (ξ', η', ζ'), indem wir Parallelstrahlen zu den Tangenten der Kurven $v =$ konst. und $u =$ konst. vom Koordinatenanfang ausgehen lassen und auf ihnen die Längen $\sqrt{u}$ bzw. $\sqrt{v}$ abtragen. Dann ist:

(7) $$\xi = \frac{2}{3}x_u, \quad \xi' = \frac{2}{3}x_v,$$

(8) $$\sum \xi^2 = u, \quad \sum \xi'^2 = v, \quad \sum \xi\xi' = -1.$$

Hieraus folgt, wenn berücksichtigt wird, daß $\xi_v = \xi'_u$ ist:

(9) $$\sum \xi'\,d\xi = 0, \quad \sum \xi\,d\xi' = 0.$$

Diese Relationen besagen im Verein mit der dritten Formel (8), daß *die beiden Hilfsflächen* $(\xi, \ldots)$ *und* $(\xi', \ldots)$ *polarreziprok bezüglich der imaginären Einheitskugel* $x^2 + y^2 + z^2 = -1$ *sind*[15]).

---

[13]) Die partiellen Ableitungen nach den Variablen u und v bzw. α und β deuten wir durch Buchstabenindizes an.

[14]) Der Kürze halber soll im folgenden der Hinweis auf das Bestehen der analogen Formeln bezüglich der beiden anderen Koordinatenachsen überall, wo ein Mißverständnis ausgeschlossen scheint, unterdrückt werden.

[15]) Die Flächen (ξ, η, ζ) und $(-\xi', -\eta', -\zeta')$ sind also polarreziprok in bezug auf die reelle Einheitskugel um den Koordinatenanfang.

Die Gleichungen (5) lassen sich jetzt folgendermaßen schreiben:

$$\xi_u = \frac{v\xi + \xi'}{2(uv-1)} + \frac{2}{3} LX, \qquad \xi_v = \frac{2}{3} MX,$$

$$\xi'_u = \frac{2}{3} MX, \qquad \xi'_v = \frac{u\xi' + \xi}{2(uv-1)} + \frac{2}{3} NX.$$

Mit Benutzung der ersten Relation (6) findet man weiter:

$$(10) \quad \begin{cases} \sum d\xi^2 = \dfrac{v\,du^2 + dv^2}{4(uv-1)} + \dfrac{4}{9} L(L\,du^2 + 2M\,du\,dv + N\,dv^2), \\ \sum d\xi'^2 = \dfrac{du^2 + u\,dv^2}{4(uv-1)} + \dfrac{4}{9} N(L\,du^2 + 2M\,du\,dv + N\,dv^2), \end{cases}$$

$$(11) \quad \sum d\xi\, d\xi' = \frac{4}{9} M(L\,du^2 + 2M\,du\,dv + N\,dv^2).$$

Der letzten dieser drei Formeln entnehmen wir eine wichtige geometrische Beziehung: *Die Asymptotenlinien der beiden Hilfsflächen* $(\xi, \ldots)$ *und* $(\xi', \ldots)$ *entsprechen stets den Asymptotenlinien der Biegungsfläche* $(x, \ldots)$.

Es soll nun gezeigt werden, daß *die beiden zueinander polarreziproken Hilfsflächen* $(\xi, \ldots)$ *und* $(\xi', \ldots)$ *der Tzitzéicaschen Klasse von Flächen angehören, bei denen das Krümmungsmaß proportional mit der 4. Potenz des Abstandes der Tangentialebene vom Koordinatenanfang ist.*

Wir beachten, daß

$$(12) \qquad -\frac{\xi'}{\sqrt{v}}, \; -\frac{\eta'}{\sqrt{v}}, \; -\frac{\zeta'}{\sqrt{v}} \quad \text{und} \quad -\frac{\xi}{\sqrt{u}}, \; -\frac{\eta}{\sqrt{u}}, \; -\frac{\zeta}{\sqrt{u}}$$

die Richtungskosinus der Flächennormalen von $(\xi, \ldots)$ und $(\xi', \ldots)$ sind, und erhalten für die Abstände w und w' der beiden Tangentialebenen vom Nullpunkt die Werte:

$$(13) \qquad w = -\sum \xi \frac{\xi'}{\sqrt{v}} = \frac{1}{\sqrt{v}}, \qquad w' = -\sum \xi' \frac{\xi}{\sqrt{u}} = \frac{1}{\sqrt{u}}.$$

Zwecks Berechnung der Krümmungsmaße k und k' bilden wir die zweiten quadratischen Fundamentalformen beider Hilfsflächen:

$$\sum d\xi\, d\left(\frac{\xi'}{\sqrt{v}}\right) = \frac{4}{9} \frac{M}{\sqrt{v}} (L\,du^2 + 2M\,du\,dv + N\,dv^2),$$

$$\sum d\xi'\, d\left(\frac{\xi}{\sqrt{u}}\right) = \frac{4}{9} \frac{M}{\sqrt{u}} (L\,du^2 + 2M\,du\,dv + N\,dv^2)$$

und finden mittels ihrer Koeffizienten und derjenigen der beiden ersten Fundamentalformen (10), wobei wieder die erste der Gleichungen (6) heranzuziehen ist, die Ausdrücke:

$$(14) \qquad k = -\frac{1}{v^2}, \qquad k' = -\frac{1}{u^2},$$

so daß sich also mit Berücksichtigung von (13) tatsächlich die Relationen

$$k = -w^4, \quad k' = -w'^4$$

ergeben.

3. Wir definieren eine weitere Fläche $(\mathfrak{x}, \mathfrak{y}, \mathfrak{z})$ durch

$$(15) \qquad \mathfrak{x} = x - \frac{2}{3} u x_u - \frac{2}{3} v x_v = x - u\xi - v\xi' \quad \text{usw.}$$

und nennen sie die zur Biegungsfläche $(x, \ldots)$ gehörige *dritte Hilfsfläche.* Sie spielt weiterhin eine wichtige Rolle beim Aufbau der Transformationstheorie. Hier sei zunächst nur bemerkt, daß ihre Punkte in den Tangentialebenen der Fläche $(x, \ldots)$ eine feste, von den Biegungen unabhängige relative Lage besitzen, also mit den Flächenelementen von $(x, \ldots)$ *starr gekoppelt* sind.

Die gleiche Bemerkung gilt für eine Schar von ∞^1 Flächen $(x^{(\nu)}, y^{(\nu)}, z^{(\nu)})$, die wir mittels der Festsetzung

$$(16) \qquad x^{(\nu)} = \mathfrak{x} + \nu\xi + \frac{1}{\nu}\xi' = x - (u - \nu)\xi - \left(v - \frac{1}{\nu}\right)\xi' \quad \text{usw.}$$

einführen, wobei ν eine willkürliche Konstante bedeutet. Diese Flächen $(x^{(\nu)}, \ldots)$ sind durch die folgenden Eigenschaften ausgezeichnet: 1. *Ihre Normalen liegen in den Tangentialebenen der Fläche* $(x, \ldots)$ *und zwar ebenfalls in starrer Koppelung; die Normalenkosinus sind den Größen* $\xi + \nu\xi'$, $\eta + \nu\eta'$, $\zeta + \nu\zeta'$ *proportional.* 2. *Zwischen den sämtlichen Flächen* $(x^{(\nu)}, \ldots)$ *und der Fläche* $(x, \ldots)$ *besteht ohne Rücksicht auf die Biegungen, denen* $(x, \ldots)$ *unterworfen wird, stets Korrespondenz der Asymptotenlinien*[16]).

Die Herleitung dieser Sätze an der Hand der aufgestellten Formeln bietet keine Schwierigkeiten, erfordert aber eine längere Rechnung, auf deren Wiedergabe wir im Hinblick auf einen späteren Beweis verzichten, der am Schluß von § 3 mit anderen Hilfsmitteln geführt werden wird.

---

[16]) In bezug auf Flächenpaare mit rechtwinklig sich kreuzenden Normalen und korrespondierenden Asymptotenlinien, für die die Fläche $(x, \ldots)$ im Verein mit einer jeden der Flächen $(x^{(\nu)}, \ldots)$ ein spezielles Beispiel liefert, vgl. man Jonas, Sur une transformation qui dépend d'une équation aux dérivées partielles du 3me ordre. C. R. de l'Ac. des Sc. **156** (1913), S. 1816. — Ohne nähere Ausführungen sei noch erwähnt, daß die Existenz der Flächen $(x^{(\nu)}, \ldots)$ aufs engste mit der Tatsache zusammenhängt daß die betrachtete Fläche $(x, \ldots)$ auf ∞^2 Weisen durch Biegung in eine tetraedrale Fläche $Ax^{2/3} + By^{2/3} + Cz^{2/3} = 1$ übergeführt werden kann. Es gibt dann immer drei Werte von ν, für die die Punkte $(x^{(\nu)}, \ldots)$ die Schnittpunkte der Tangentialebenen mit den Koordinatenachsen werden.

§ 2.

Einführung der Parameter der Asymptotenlinien; Konstruktion der Biegungsfläche auf Grund einer Lösung der partiellen Differentialgleichung $\theta_{\alpha\beta} = e^{\theta} - e^{-2\theta}$.

1. Wir führen nun die Parameter α und β der Asymptotenlinien ein und richten unser Augenmerk besonders auf die quadratische Differentialform § 1 (11), die, auf α und β bezogen, sich folgendermaßen schreiben läßt:

(1) $$\sum d\xi\, d\xi' = 2\, e^{\theta}\, d\alpha\, d\beta.$$

Es ist bekannt, daß die *normierten*, d. h. durch die 4. Wurzel aus dem absoluten Betrag des Krümmungsmaßes dividierten Normalenkosinus einer Fläche als Funktionen der Parameter der Asymptotenlinien eine partielle Differentialgleichung 2. Ordnung vom Moutardschen Typus erfüllen. Wir wenden diesen Satz zunächst auf die Hilfsflächen $(\xi, \ldots)$ und $(\xi', \ldots)$ an, deren Asymptotenlinien, wie wir sahen, denen der Biegungsfläche $(x, \ldots)$ entsprechen, und schließen mit Rücksicht auf § 1 (12) und (14), daß eben diese Größen ξ, η, ζ und ξ', η', ζ' Integrale zweier Moutardscher Gleichungen

(2) $$\xi_{\alpha\beta} = \mathsf{M}\xi, \qquad \xi'_{\alpha\beta} = \mathsf{M}'\xi'$$

sind. Nach § 1 (8) und (9) ist aber:

(3) $$\sum \xi\xi' = -1, \quad \sum \xi_\alpha \xi' = 0, \quad \sum \xi_\beta \xi' = 0, \quad \sum \xi'_\alpha \xi = 0, \quad \sum \xi'_\beta \xi = 0.$$

Differentiiert man die zweite und die vierte dieser Gleichungen nach β, die dritte und die fünfte nach α und macht Gebrauch von (1) und (2), so findet man:

(4) $$\mathsf{M} = \mathsf{M}' = \sum \xi_\alpha \xi'_\beta = \sum \xi_\beta \xi'_\alpha = e^{\theta},$$

während außerdem infolge von (1)

(5) $$\sum \xi_\alpha \xi'_\alpha = 0, \quad \sum \xi_\beta \xi'_\beta = 0$$

ist. Da aber ξ, η, ζ (Entsprechendes gilt von ξ', η', ζ') nicht nur normierte Normalenkosinus sind, sondern gleichzeitig die laufenden Koordinaten einer auf die Asymptotenlinien (α, β) bezogenen Fläche darstellen, so bestehen neben (2) je zwei weitere Reihen von Relationen von der Form:

$$\xi_{\alpha\alpha} = a\xi_\alpha + p\xi_\beta, \qquad \xi_{\beta\beta} = q\xi_\alpha + b\xi_\beta.$$

Um zunächst den Koeffizienten a zu bestimmen, multiplizieren wir die erste dieser beiden Gleichungen mit ξ'_β und addieren die beiden analogen Beziehungen. Unter Berücksichtigung von (4) ergibt sich:

$$\sum \xi_{\alpha\alpha}\xi'_\beta = a e^{\theta}.$$

Differentiiert man andererseits die Gleichung $\sum \xi_\alpha \xi'_\beta = e^\theta$ nach α, so findet man:

$$\sum \xi_{\alpha\alpha} \xi'_\beta = e^\theta \theta_\alpha .$$

Es wird also $a = \theta_\alpha$ und entsprechend $b = \theta_\beta$. Mithin haben wir für ξ, η, ζ das folgende System von Differentialgleichungen:

$$(6) \qquad \xi_{\alpha\alpha} = \theta_\alpha \xi_\alpha + p \xi_\beta, \quad \xi_{\alpha\beta} = e^\theta \xi, \quad \xi_{\beta\beta} = q \xi_\alpha + \theta_\beta \xi_\beta,$$

für das man als Integrabilitätsbedingungen die Relationen

$$(7) \qquad p_\beta + p\theta_\beta = 0, \quad q_\alpha + q\theta_\alpha = 0, \quad \theta_{\alpha\beta} + pq = e^\theta$$

erhält.

Wir wollen nun von der spezialisierenden Annahme absehen, daß eine der Größen p und q oder daß alle beide identisch verschwinden, wodurch sich die dritte der Relationen (7) auf die Liouvillesche Differentialgleichung für θ reduzieren würde. Das bedeutet also, daß die Hilfsfläche $(\xi, \ldots)$ nicht Regelfläche oder gar Fläche 2. Grades sein soll. Aus der Gesamtheit der Biegungsflächen vom Typus § 1 (1) scheiden damit als nicht ohne weiteres der im folgenden entwickelten Transformationstheorie zugänglich die in der Einleitung erwähnten tetraedralen Flächen aus, für die die beiden Hilfsflächen Flächen 2. Grades werden. Unter dieser die Allgemeinheit nicht beschränkenden, sondern nur die Ausartungsfälle ausschließenden Voraussetzung folgt aus (7):

$$(8) \qquad \begin{gathered} p = \varphi(\alpha) e^{-\theta}, \quad q = \psi(\beta) e^{-\theta}, \\ \theta_{\alpha\beta} + \varphi(\alpha)\psi(\beta) e^{-2\theta} = e^\theta. \end{gathered}$$

An Stelle von α und β führen wir jetzt die Größen

$$\bar{\alpha} = \int \sqrt[3]{\varphi(\alpha)}\, d\alpha, \quad \bar{\beta} = \int \sqrt[3]{\psi(\beta)}\, d\beta$$

als neue Variablen ein und setzen außerdem

$$e^\theta = e^{\bar{\theta}} \sqrt[3]{\varphi(\alpha)} \sqrt[3]{\psi(\beta)} .$$

Dabei behält die Differentialform (1) ihre Gestalt:

$$\sum d\xi\, d\xi' = 2 e^\theta d\alpha\, d\beta = 2 e^{\bar{\theta}} d\bar{\alpha}\, d\bar{\beta},$$

während (8) in

$$\bar{\theta}_{\bar{\alpha}\bar{\beta}} = e^{\bar{\theta}} - e^{-2\bar{\theta}}$$

übergeht. Gleichzeitig vereinfachen sich die Koeffizienten im Gleichungssystem (6). Schreiben wir nachträglich wieder α, β, θ für $\bar{\alpha}, \bar{\beta}, \bar{\theta}$, so lauten die Formeln:

$$(9) \qquad \theta_{\alpha\beta} = e^\theta - e^{-2\theta},$$

$$(10) \qquad \xi_{\alpha\alpha} = \theta_\alpha \xi_\alpha + e^{-\theta} \xi_\beta, \quad \xi_{\alpha\beta} = e^\theta \xi, \quad \xi_{\beta\beta} = e^{-\theta} \xi_\alpha + \theta_\beta \xi_\beta .$$

Dies sind die von Tzitzéica[17]) aufgestellten Differentialgleichungen, von deren Integration die Bestimmung der nicht-geradlinigen Flächen mit der Eigenschaft $\frac{K}{W^4} = \text{konst.}$ abhängt. Bei Tzitzéica steht h an der Stelle von e^θ. Die Schreibung mittels der Exponentialfunktion soll negative Werte von e^θ, bei denen θ die imaginäre additive Konstante $i\pi$ aufnimmt, nicht ausschließen, schien aber aus formalen Gründen vorteilhaft.

Wird ein Integral θ von (9) als bekannt vorausgesetzt, so ist (10) als ein unbeschränkt integrabeles System simultaner totaler Differentialgleichungen für $\xi, \xi_\alpha, \xi_\beta$ aufzufassen. Die Integration desselben führt drei willkürliche Konstanten ein, nämlich die Anfangswerte von $\xi, \xi_\alpha, \xi_\beta$ für ein spezielles Wertepaar $\alpha = \alpha_0$, $\beta = \beta_0$ der Parameter. Sind drei solcher Tripel von Anfangswerten so gewählt, daß

$$\begin{vmatrix} \xi, & \xi_\alpha, & \xi_\beta \\ \eta, & \eta_\alpha, & \eta_\beta \\ \zeta, & \zeta_\alpha, & \zeta_\beta \end{vmatrix}_{\substack{\alpha=\alpha_0 \\ \beta=\beta_0}} \neq 0$$

ist, so stellen ξ, η, ζ drei voneinander linear-unabhängige Integrale des Systems (10) dar, während jedes weitere Integral sich mittels konstanter Koeffizienten linear durch ξ, η, ζ ausdrücken läßt.

2. Es sei also ξ, η, ζ ein solches Tripel linear-unabhängiger Lösungen. Wir setzen

$$(\xi, \xi_\alpha, \xi_\beta) = \begin{vmatrix} \xi, & \xi_\alpha, & \xi_\beta \\ \eta, & \eta_\alpha, & \eta_\beta \\ \zeta, & \zeta_\alpha, & \zeta_\beta \end{vmatrix} = \varDelta$$

und finden mit Hilfe von (10):

$$\frac{\partial \log \varDelta}{\partial \alpha} = \theta_\alpha, \qquad \frac{\partial \log \varDelta}{\partial \beta} = \theta_\beta,$$

so daß also $\varDelta = c e^\theta$ ($c = \text{konst.}$) folgt. Die Größen ξ, η, ζ mögen nun, was infolge der Homogenität der Differentialgleichungen (10) ohne weiteres möglich ist, durch Hinzufügung konstanter Faktoren so *normiert* werden, daß $c = -1$, also

(11) $$\varDelta = -e^\theta$$

wird. Aus den Gleichungen

$$\sum \xi' \xi = -1, \qquad \sum \xi' \xi_\alpha = 0, \qquad \sum \xi' \xi_\beta = 0$$

ergibt sich dann:

(12) $$\xi' = e^{-\theta}(\eta_\alpha \zeta_\beta - \zeta_\alpha \eta_\beta) \quad \text{usw.}^{18)}$$

---

[17]) S. die Zitate in der Einleitung, besonders unter [5]).

[18]) Die beiden anderen Formeln erhält man stets durch zyklische Vertauschung.

und hieraus, indem man beim Differentiieren von (10) Gebrauch macht:

(13) $$\xi'_\alpha = -(\eta\zeta_\alpha - \zeta\eta_\alpha), \quad \xi'_\beta = \eta\zeta_\beta - \zeta\eta_\beta.$$

Durch abermalige Differentiation gewinnt man das zu (10) duale System:

(14) $$\xi'_{\alpha\alpha} = \theta_\alpha \xi'_\alpha - e^{-\theta}\xi'_\beta, \quad \xi'_{\alpha\beta} = e^\theta \xi', \quad \xi'_{\beta\beta} = -e^{-\theta}\xi'_\alpha + \theta_\beta \xi'_\beta.$$

Man bestätigt ferner die Beziehungen:

$$\sum \xi\xi'_\alpha = 0, \quad \sum \xi\xi'_\beta = 0, \quad \sum \xi_\alpha \xi'_\alpha = 0, \quad \sum \xi_\beta \xi'_\beta = 0,$$
$$\sum \xi_\alpha \xi'_\beta = \sum \xi_\beta \xi'_\alpha = e^\theta$$

und erhält als Gegenstücke zu (11), (12) und (13) die Formeln:

(15) $$\Delta' = (\xi', \xi'_\alpha, \xi'_\beta) = -e^\theta,$$

(16) $$\xi = e^{-\theta}(\eta'_\alpha \zeta'_\beta - \zeta'_\alpha \eta'_\beta),$$

(17) $$\xi_\alpha = -(\eta'\zeta'_\alpha - \zeta'\eta'_\alpha), \quad \xi_\beta = \eta'\zeta'_\beta - \zeta'\eta'_\beta.$$

Wir beachten, daß die Relationen (17) und (13) für die Flächen $(\xi, \ldots)$ und $(\xi', \ldots)$ die Darstellung mittels der Lelieuvreschen Formeln bedeuten, und schließen unmittelbar aus dieser Tatsache, daß die Krümmungsmaße durch

$$k = -\frac{1}{v^2}, \quad k' = -\frac{1}{u^2}$$

gegeben sind, wobei

(18) $$\sum \xi^2 = u, \quad \sum \xi'^2 = v$$

gesetzt ist. Man verifiziert leicht die Zugehörigkeit der Flächen zur Tzitzéicaschen Klasse, nämlich das Bestehen der Relationen $k = -w^4$, $k' = -w'^4$.

3. Nach diesen Vorbereitungen erhalten wir die laufenden Koordinaten der gesuchten Biegungsfläche (x, y, z) durch drei Quadraturen:

(19) $$x = \frac{3}{2}\int(\xi\,du + \xi'\,dv) \quad \text{usw.}$$

In der Tat ergibt sich als Quadrat des Linienelements sofort der Ausdruck § 1 (1). Es bedarf nur noch des Nachweises, daß in (19) unter dem Integralzeichen ein exaktes Differential steht, d. h. daß

$$\Omega = \xi_\beta u_\alpha + \xi'_\beta v_\alpha - \xi_\alpha u_\beta - \xi'_\alpha v_\beta = 0$$

wird. Infolge von (18) kann man aber schreiben:

$$\frac{1}{2}\Omega = -\eta(\xi_\alpha\eta_\beta - \eta_\alpha\xi_\beta) + \zeta(\zeta_\alpha\xi_\beta - \xi_\alpha\zeta_\beta) - \eta'(\xi'_\alpha\eta'_\beta - \eta'_\alpha\xi'_\beta)$$
$$+ \zeta'(\zeta'_\alpha\xi'_\beta - \xi'_\alpha\zeta'_\beta)$$

und findet mit Berücksichtigung von (12) und (16) $\Omega = 0$.

Wir fassen das gewonnene Ergebnis zusammen:

Um eine (nicht spezielle) Biegungsfläche vom Typus

$$\sum dx^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2)$$

zu konstruieren, hat man von einem Integral θ der partiellen Differentialgleichung

$$\theta_{\alpha\beta} = e^{\theta} - e^{-2\theta}$$

auszugehen, ein Tripel ξ, η, ζ von linear unabhängigen und nach geeigneter Normierung die Bedingung $\Delta = (\xi, \xi_\alpha, \xi_\beta) = -e^{\theta}$ erfüllenden Lösungen des Systems

$$\xi_{\alpha\alpha} = \theta_\alpha \xi_\alpha + e^{-\theta}\xi_\beta, \qquad \xi_{\alpha\beta} = e^{\theta}\xi, \qquad \xi_{\beta\beta} = e^{-\theta}\xi_\alpha + \theta_\beta \xi_\beta$$

aufzusuchen und

$$\xi' = e^{-\theta}(\eta_\alpha\zeta_\beta - \zeta_\alpha\eta_\beta), \quad \eta' = e^{-\theta}(\zeta_\alpha\xi_\beta - \xi_\alpha\zeta_\beta), \quad \zeta' = e^{-\theta}(\xi_\alpha\eta_\beta - \eta_\alpha\xi_\beta),$$

$$u = \xi^2 + \eta^2 + \zeta^2, \qquad v = \xi'^2 + \eta'^2 + \zeta'^2$$

zu bilden. Alsdann sind

$$x = \frac{3}{2}\int(\xi\,du + \xi'\,dv), \quad y = \frac{3}{2}\int(\eta\,du + \eta'\,dv), \quad z = \frac{3}{2}\int(\zeta\,du + \zeta'\,dv)$$

die laufenden Koordinaten der Biegungsfläche.

4. Wir fügen noch eine Bemerkung in bezug auf die beiden zueinander dualen Systeme (10) und (14) hinzu. Ist ϑ ein Integral von (10), ϑ' ein beliebiges Integral von (14), so besteht, wovon man sich leicht durch Differentiieren überzeugt, eine Relation von der Form:

$$e^{-\theta}(\vartheta_\alpha\vartheta'_\beta + \vartheta_\beta\vartheta'_\alpha) - \vartheta\vartheta' = \text{konst.}$$

Insbesondere ist

$$e^{-\theta}(\xi_\alpha\xi'_\beta + \xi_\beta\xi'_\alpha) = \xi\xi' - 1$$

usw., während für zwei verschiedene Buchstaben, z. B. η und ζ', die Konstante verschwindet, also

$$(20) \qquad e^{-\theta}(\eta_\alpha\zeta'_\beta + \eta_\beta\zeta'_\alpha) - \eta\zeta' = 0$$

wird. Man bestätigt diese Beziehungen an der Hand von (12) und (13) durch Einsetzen der Ausdrücke für die gestrichenen Größen.

§ 3.

Satz über die Bestimmung der Asymptotenlinien auf den betrachteten Biegungsflächen; Vorbereitung der Transformationstheorie mittels der Lelieuvreschen Relationen.

1. Die Betrachtung der quadratischen Differentialform

$$(1) \qquad \sum d\xi\,d\xi' = \frac{4}{9}M(L\,du^2 + 2M\,du\,dv + N\,dv^2) = 2e^{\theta}\,d\alpha\,d\beta$$

führt zur Auffindung einer bemerkenswerten Eigenschaft, die die Flächen der in Rede stehenden isometrischen Klasse mit den auf die Flächen 2. Grades abwickelbaren Flächen teilen. Es gilt der Satz:

Auf den Biegungsflächen vom Typus

$$ds^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2)$$

lassen sich die Asymptotenlinien mit Hilfe von Quadraturen bestimmen.

Es sei nämlich K das Krümmungsmaß der Form (1), das, wenn eine durch L, M, N definierte Biegungsfläche vorliegt, als Funktion der Parameter u und v dargestellt werden kann. Andererseits erhält man auf Grund der Differentialgleichung § 2 (9):

$$\mathsf{K} = -e^{-\theta}\theta_{\alpha\beta} = -1 + e^{-3\theta}. \quad {}^{19})$$

Es bietet sich demnach in

$$e^{-\theta} = \sqrt[3]{1 + \mathsf{K}}$$

ein Multiplikator, durch den (1) in das Produkt zweier Differentiale, also in eine Form von der Krümmung 0 übergeht. Es ist aber bekannt, daß bei einer quadratischen Differentialform von verschwindender Krümmung die Zerlegung in das Produkt zweier Differentiale durch Quadraturen gelingt; und zwar sind im allgemeinen *drei* Quadraturen erforderlich, eine zur Bestimmung der beiden zueinander reziproken integrierenden Faktoren und zwei weitere zur Bestimmung der neuen Variablen.

Der hiermit bewiesene Satz liegt übrigens wesentlich tiefer als der gleichlautende für die Biegungsflächen der Flächen 2. Grades. Dort gelingt es nämlich, bei Beziehung der Fläche auf die verbogenen Erzeugenden die integrierenden Faktoren ohne Quadratur durch die zweiten Fundamentalgrößen L, M, N in geschlossener Form auszudrücken[20]), während man im vorliegenden Falle bereits bei dem Versuch, den Beweis unmittelbar auf die Gestalt des Linienelements und auf die von L, M, N erfüllten Gleichungen (6) von § 1 zu stützen, auf außerordentliche rechnerische Schwierigkeiten stößt.

Wir erinnern noch daran, daß nach Beltrami[21]) der Multiplikator λ, durch den eine quadratische Differentialform in das Produkt zweier Differentiale verwandelt wird, der partiellen Differentialgleichung 2. Ordnung

$$\frac{1}{2}\Delta_2 \log \lambda = \mathsf{K}$$

---

[19]) Bekanntlich ist das Krümmungsmaß einer auf Minimalparameter bezogenen quadratischen Differentialform $2f\,d\alpha\,d\beta$ durch die Formel: $\mathsf{K} = -\frac{1}{f}\frac{\partial^2 \log f}{\partial\alpha\,\partial\beta}$ gegeben.

[20]) Jonas, Über eine neue partielle Differentialgleichung des Deformationsproblems usw. Berl. Math. Ges. Ber. **13** (1914), S. 52.

[21]) Siehe z. B. Knoblauch, Grundlagen der Differentialgeometrie (1913), S. 581.

genügt, in der K das Krümmungsmaß der gegebenen Form und der Operator Δ_2 den 2. Differentialparameter bedeutet. Danach besteht für die Form (1) die Relation:

$$\frac{1}{6}\Delta_2 \log(1+\mathsf{K}) = \mathsf{K},$$

durch die die Differentialgleichung

$$\theta_{\alpha\beta} = e^{\theta} - e^{-2\theta}$$

unabhängig von der Wahl der Bezugsparameter dargestellt wird.

2. An der Hand der Entwicklungen von § 2 gelingt nun auch die Aufstellung der Lelieuvreschen Formeln und der Moutardschen Gleichung für die Biegungsfläche $(x, \ldots)$. Da für diese nach § 1 (4)

$$K = -\frac{1}{9(uv-1)^2}$$

ist, sind die normierten Normalenkosinus:

$$\frac{1}{\sqrt[4]{-K}} X = \sqrt{3}(\eta\zeta' - \zeta\eta') \quad \text{usw.}$$

Wir ziehen es vor, unter Fortlassung des Faktors $\sqrt{3}$ die Größen

$$(2) \qquad \Xi = \eta\zeta' - \zeta\eta', \qquad \mathsf{H} = \zeta\xi' - \xi\zeta', \qquad \mathsf{Z} = \xi\eta' - \eta\xi'$$

zu benutzen, was nur zur Folge hat, daß die sich ergebenden Ausdrücke für x_α und x_β sich von der gewöhnlichen Gestalt der Lelieuvreschen Formeln durch das Auftreten des Faktors 3 unterscheiden.

Man findet mittels einer leichten Umformung:

$$-(\mathsf{H}\mathsf{Z}_\alpha - \mathsf{Z}\mathsf{H}_\alpha) = -\xi\sum\xi(\eta'\zeta'_\alpha - \zeta'\eta'_\alpha) - \xi'\sum\xi'(\eta\zeta_\alpha - \zeta\eta_\alpha);$$

die rechte Seite geht mit Benutzung von § 2 (13) und (17) in

$$\xi\sum\xi\xi_\alpha + \xi'\sum\xi'\xi'_\alpha = \frac{1}{2}u_\alpha\xi + \frac{1}{2}v_\alpha\xi'$$

über. Wir erhalten also, wobei wir den entsprechenden Ausdruck für die Ableitung nach β hinzufügen:

$$(3) \qquad \begin{cases} x_\alpha = \frac{3}{2}(u_\alpha\xi + v_\alpha\xi') = -3(\mathsf{H}\mathsf{Z}_\alpha - \mathsf{Z}\mathsf{H}_\alpha), \\ x_\beta = \frac{3}{2}(u_\beta\xi + v_\beta\xi') = 3(\mathsf{H}\mathsf{Z}_\beta - \mathsf{Z}\mathsf{H}_\beta). \end{cases}$$

Die erste der Formeln (2) werde nun nacheinander nach α und β differentiiert. Berücksichtigt man, daß $\xi, \eta, \zeta, \xi', \eta', \zeta'$ die Differentialgleichung $\vartheta_{\alpha\beta} = e^{\theta}\vartheta$ befriedigen und daß infolge von § 2 (20)

$$\eta_\alpha\zeta'_\beta + \eta_\beta\zeta'_\alpha - \eta'_\beta\zeta_\alpha - \eta'_\alpha\zeta_\beta = e^{\theta}(\eta\zeta' - \zeta\eta') = e^{\theta}\Xi$$

ist, so ergibt sich die von Ξ, H, Z erfüllte Moutardsche Differentialgleichung:

(4) $$\Xi_{\alpha\beta} = 3e^{\theta}\,\Xi.$$

Von ihr hängt die infinitesimale Verbiegung der Fläche $(x, \ldots)$ ab sowie die Konstruktion der *W*-Kongruenzen, für die $(x, \ldots)$ den einen Brennflächenmantel bildet. Wir werden von dieser Bemerkung später Gebrauch machen.

Zur Vereinfachung der Schreibweise in weiterhin folgenden Entwickelungen, die sich an die Moutardsche Transformation knüpfen, führen wir ein abkürzendes Symbol ein. Es sei allgemein, wenn φ und ψ zwei Integrale einer Moutardschen Gleichung sind,

$$\int[-(\varphi\psi_\alpha - \psi\varphi_\alpha)d\alpha + (\varphi\psi_\beta - \psi\varphi_\beta)d\beta] = \langle\varphi, \psi\rangle$$

gesetzt. Wir schreiben danach (3) in der Form:

$$x = 3\langle \mathrm{H}, \mathrm{Z}\rangle, \qquad y = 3\langle \mathrm{Z}, \Xi\rangle, \qquad z = 3\langle \Xi, \mathrm{H}\rangle.$$

3. Wir betrachten ferner die in § 1, **3** eingeführte *dritte Hilfsfläche* $(\mathfrak{x}, \mathfrak{y}, \mathfrak{z})$, gegeben durch

$$\mathfrak{x} = x - u\xi - v\xi' \quad \text{usw.}$$

Differentiiert man $\mathfrak{x}$ nach α und beachtet dabei, daß $u = \sum \xi^2$, $v = \sum \xi'^2$ ist, so ergibt sich zunächst:

$$\mathfrak{x}_\alpha = \eta(\xi\eta_\alpha - \eta\xi_\alpha) - \zeta(\zeta\xi_\alpha - \xi\zeta_\alpha) + \eta'(\xi'\eta'_\alpha - \eta'\xi'_\alpha) - \zeta'(\zeta'\xi'_\alpha - \xi'\zeta'_\alpha).$$

Die weitere Umformung gelingt wieder mittels der Relationen (13) und (17) von § 2. Analog ist $\mathfrak{x}_\beta$ zu bilden. Es wird:

$$\begin{aligned} \mathfrak{x}_\alpha &= -(\eta\zeta'_\alpha - \zeta'\eta_\alpha + \eta'\zeta_\alpha - \zeta\eta'_\alpha), \\ \mathfrak{x}_\beta &= \quad \eta\zeta'_\beta - \zeta'\eta_\beta + \eta'\zeta_\beta - \zeta\eta'_\beta \end{aligned}$$

oder, wenn wir uns der eben eingeführten Bezeichnung bedienen:

(5) $$\mathfrak{x} = \langle\eta, \zeta'\rangle + \langle\eta', \zeta\rangle \quad \text{usw.}$$

Die Aufstellung dieser Formelgruppe ist einer der wichtigsten vorbereitenden Schritte für unsere Zwecke. Wir sind damit im Besitze eines Systems *erweiterter* Lelieuvrescher Formeln, das bei einer Moutardschen Transformation der sechs Größen $\xi, \eta, \zeta, \xi', \eta', \zeta'$ die gleiche Verwendung gestattet wie die eigentlichen Lelieuvreschen Formeln bei der Konstruktion der *W*-Kongruenzen.

4. Im Anschluß an (5) und mit besonderer Berücksichtigung der Relationen § 2 (13) und (17) findet man für eine jede der in § 1, **3** betrachteten und durch

(6) $$x^{(\nu)} = \mathfrak{x} + \nu\xi + \frac{1}{\nu}\xi' = x - (u - \nu)\xi - \left(v - \frac{1}{\nu}\right)\xi'$$

definierten Flächen $(x^{(\nu)}, y^{(\nu)}, z^{(\nu)})$ die Darstellung:

$$x^{(\nu)} = \frac{1}{\nu} \langle \eta + \nu\eta', \zeta + \nu\zeta' \rangle \quad \text{usw.} \tag{7}$$

Das sind aber die Lelieuvreschen Formeln für eine auf ihre Asymptotenlinien bezogene Fläche, deren Normalenkosinus sich wie $\xi + \nu\xi' : \eta + \nu\eta' : \zeta + \nu\zeta'$ verhalten. Damit sind die am Schluß von § 1 ausgesprochenen Eigenschaften der Flächen $(x^{(\nu)}, \ldots)$ bewiesen.

Erwähnt sei noch eine weitere Eigentümlichkeit dieser Flächen: *Das Krümmungsmaß einer Fläche* $(x^{(\nu)}, \ldots)$ *ändert sich nicht, wenn die Fläche* $(x, \ldots)$ *verbogen wird.*

Es ergibt sich nämlich auf Grund von (7) der folgende, u und v nur explizite enthaltende Ausdruck für $K^{(\nu)}$:

$$K^{(\nu)} = -\frac{\nu^2}{[\Sigma(\xi + \nu\xi')^2]^2} = -\frac{1}{\left(\frac{1}{\nu} u + \nu v - 2\right)^2}.$$

§ 4.

Anwendung der inhaltstreuen Affinität auf die Tzitzéicaschen Hilfsflächen und Gegenstück zur Lieschen Transformation der Flächen von konstanter Krümmung.

1. Ehe wir an unsere Hauptaufgabe, die Aufstellung einer Transformationstheorie für die Flächen der betrachteten isometrischen Klasse, herantreten, mögen noch zwei Eigenschaften der zu ihrer Bestimmung dienenden Differentialgleichungen zur Sprache gebracht werden, die ihrerseits übrigens auch schon die Konstruktion neuer Biegungsflächen aus einer gegebenen gestatten. Es ist das einmal die leicht ersichtliche Tatsache, daß das Lösungstripel ξ, η, ζ des Systems § 2 (10) einer homogenen linearen Transformation unterworfen werden kann, deren Determinante mit Rücksicht auf § 2 (11) den Wert 1 haben muß. Geometrisch bedeutet dies eine inhaltstreue affine Raumtransformation mit dem Zentrum im Koordinatenanfang. Da nun aber eine Drehung der Hilfsfläche $(\xi, \ldots)$ keine neue Biegungsfläche entstehen läßt[22]), haben wir eine von Drehung freie Affinität zu nehmen, die von fünf willkürlichen Konstanten abhängt.

Wir wollen zeigen, daß die Darstellung der zugehörigen ∞^5 neuen Biegungsflächen im allgemeinen von drei Quadraturen abhängt. Dabei mögen die Koordinatenachsen mit den Hauptdilatationsachsen zusammenfallen, was zur Folge hat, daß drei der verfügbaren Konstanten nicht in

---

[22]) Man beachte die Invarianz der Ausdrücke $u = \Sigma\xi^2$ und $v = \Sigma\xi'^2$ den Drehungen gegenüber.

Erscheinung treten. An Stelle von ξ, η, ζ ist $c_1\xi, c_2\eta, c_3\zeta$ zu schreiben, wobei die Konstanten c_1, c_2, c_3 die Relation

$$(1) \qquad c_1 c_2 c_3 = 1$$

erfüllen müssen. Man erkennt unschwer, daß ξ', η', ζ' durch $\frac{1}{c_1}\xi', \frac{1}{c_2}\eta', \frac{1}{c_3}\zeta'$ zu ersetzen sind. Wird

$$c_1^2\xi^2 + c_2^2\eta^2 + c_3^2\zeta^2 = \tilde{u}, \qquad \frac{1}{c_1^2}\xi'^2 + \frac{1}{c_2^2}\eta'^2 + \frac{1}{c_3^2}\zeta'^2 = \tilde{v}$$

gesetzt, so ist die neue Biegungsfläche $(\tilde{x}, \tilde{y}, \tilde{z})$ durch

$$\tilde{x} = \frac{3}{2}\int\left(c_1^2\xi\, d\tilde{u} + \frac{1}{c_1^2}\xi' d\tilde{v}\right) \quad \text{usw.}$$

bestimmt. Führt man die Integration teilweise aus und beachtet (1), so findet man:

$$(2) \qquad \tilde{x} = c_1^3\xi^3 + \frac{1}{c_1^3}\xi'^3 + 3\frac{c_2}{c_3}\int(\xi\eta\, d\eta + \xi'\zeta' d\zeta') + 3\frac{c_3}{c_2}\int(\xi\zeta\, d\zeta + \xi'\eta' d\eta')$$

usw.; $\tilde{y}$ und $\tilde{z}$ folgen durch zyklische Permutation innerhalb der Tripel $\xi\eta\zeta$, $\xi'\eta'\zeta'$ und $c_1c_2c_3$. Daß unter den Integralzeichen exakte Differentiale stehen, bestätigt man auch ohne Benutzung von α und β an der Hand der Formeln von § 1. Die Anzahl der durch (2) geforderten Quadraturen reduziert sich (c_1, c_2, c_3 als verschieden vorausgesetzt) auf drei, da die Summen

$$3\int(\xi\eta\, d\eta + \xi'\zeta' d\zeta') + 3\int(\xi\zeta\, d\zeta + \xi'\eta' d\eta') = x - \xi^3 - \xi'^3 \quad \text{usw.}$$

bekannt sind.

Wir beachten noch, daß an die Stelle von Ξ, H, Z die Größen

$$\frac{c_2}{c_3}\eta\zeta' - \frac{c_3}{c_2}\zeta\eta', \qquad \frac{c_3}{c_1}\zeta\xi' - \frac{c_1}{c_3}\xi\zeta', \qquad \frac{c_1}{c_2}\xi\eta' - \frac{c_2}{c_1}\eta\xi'$$

treten, und schließen hieraus auf die auch leicht direkt erweisbare Tatsache, daß die Produkte $\eta\zeta'$, $\zeta\eta'$ usw. einzeln der Moutardschen Gleichung § 3 (4):

$$(3) \qquad \Theta_{\alpha\beta} = 3e^\theta\,\Theta$$

genügen. Allgemein sei bemerkt, daß, wenn ϑ und ϑ' zwei Integrale der Systeme § 2 (10) und (14) darstellen, so daß nach § 2, 4 gleichzeitig die Beziehung

$$(4) \qquad e^{-\theta}(\vartheta_\alpha\vartheta'_\beta + \vartheta_\beta\vartheta'_\alpha) - \vartheta\vartheta' = 3C \qquad (C = \text{konst.})$$

besteht, der Ausdruck $\vartheta\vartheta' + C$ ein Integral von (3) ist.

2. Die zweite Bemerkung betrifft die Tatsache, daß die Differentialgleichung

$$(5) \qquad \theta_{\alpha\beta} = e^\theta - e^{-2\theta}$$

wie jede Gleichung der Form $\theta_{\alpha\beta} = f(\theta)$ neben einem gegebenen Integral $\theta(\alpha, \beta)$ noch ∞^1 Lösungen $\theta\left(c\alpha, \frac{\beta}{c}\right)$ zuläßt, wobei c eine willkürliche Konstante bedeutet. Man erkennt, daß es sich hier um ein *Gegenstück zur Lieschen Transformation der pseudosphärischen Flächen* handelt. Die Einführung des neuen Integrals $\theta\left(c\alpha, \frac{\beta}{c}\right)$ verändert das System der Differentialgleichungen § 2 (10), das in der neuen Gestalt erst wieder integriert werden muß, so daß, abgesehen von der Kenntnis des Integrals θ, sich keine Vereinfachung erzielen läßt. Schreiben wir α und β an Stelle von $c\alpha$ und $\frac{\beta}{c}$, so behält die Funktion $\theta(\alpha, \beta)$ die ursprüngliche Form, während das System § 2 (10), von dem die Bestimmung der neuen ξ, η, ζ, also der Koordinaten einer, wie wir sagen können, *Lieschen Transformierten der Tzitzéicaschen Fläche* abhängt, in:

$$(6) \qquad \vartheta_{\alpha\alpha} = \theta_\alpha \vartheta_\alpha + \varkappa e^{-\theta} \vartheta_\beta, \qquad \vartheta_{\alpha\beta} = e^{\theta} \vartheta, \qquad \vartheta_{\beta\beta} = \frac{1}{\varkappa} e^{-\theta} \vartheta_\alpha + \theta_\beta \vartheta_\beta$$

übergeht, wobei $\varkappa = c^3$ ist. Gleichzeitig lautet das duale System:

$$(7) \qquad \vartheta'_{\alpha\alpha} = \theta_\alpha \vartheta'_\alpha - \varkappa e^{-\theta} \vartheta'_\beta, \qquad \vartheta'_{\alpha\beta} = e^{\theta} \vartheta', \qquad \vartheta'_{\beta\beta} = -\frac{1}{\varkappa} e^{-\theta} \vartheta'_\alpha + \theta_\beta \vartheta'_\beta.$$

Hiernach ist es, ohne daß es der bestätigenden Rechnung bedarf, klar, daß ebenso wie § 2 (10) und (14) auch die vorliegenden allgemeineren Systeme von Differentialgleichungen unbeschränkt integrabel sind, daß zwei Lösungen ϑ und ϑ' von (6) und (7) wieder die Relation (4) befriedigen und daß der aus ihnen gebildete Ausdruck $\vartheta\vartheta' + C$ wieder ein Integral von (3) ist.

Das System (6) ist von fundamentaler Bedeutung. Es beherrscht nämlich, wenn $\varkappa = \frac{1+n}{1-n}$ gesetzt wird, die Tzitzéicasche Transformation T_n der durch $\frac{K}{W^4} =$ konst. charakterisierten Flächen und gleichzeitig unsere in analytischer Beziehung damit äquivalente Transformation Θ_n, die im nächsten Paragraphen für die Biegungsflächen der betrachteten Klasse aufgestellt werden soll. Das weiterhin mit R bezeichnete Integral von (6), dem die Rolle einer *transformierenden Funktion* zufallen wird, ist auf Grund der letzten, das Analogon zur Lieschen Transformation betreffenden Bemerkung einer geometrischen Interpretation fähig, die wir hier ohne die Absicht späterer Verwendung voraufschicken: *R läßt sich als Koordinate (Abstand des laufenden Punktes von einer festen Ebene) für eine neue Fläche der Tzitzéicaschen Klasse auffassen, die aus (ξ, η, ζ) durch den zur Lieschen Transformation analogen Prozeß hervorgeht.*

Auffallend ist die Tatsache, daß eine solche Verwandtschaft zwischen zwei Transformationen der gleichen Klasse von Biegungsflächen, von denen die eine, hier die Transformation Θ_n, in Gestalt einer geometrischen Konstruktion auftritt, während die andere, hier also das Gegenstück zur Lieschen Transformation, einen ausschließlich analytischen Charakter besitzt, sich in gleicher Weise bereits bei den Biegungsflächen der Pseudosphäre findet. In der Tat läßt sich — ein darauf bezüglicher Hinweis scheint übrigens in der Literatur zu fehlen — auf die Riccatische Gleichung, von der die Bäcklundsche Transformation einer speudosphärischen Fläche abhängt, auch die Darstellung der Koordinaten einer allerdings imaginären Lieschen Transformierten zurückführen.

§ 5.

Aufstellung der Transformation Θ_n für die betrachtete Klasse von Biegungsflächen.

1. Es sei wieder eine auf ihre Asymptotenlinien bezogene und an der Hand der Entwicklungen von § 2 konstruierte Biegungsfläche (x, y, z) vom Typus

$$\sum dx^2 = \frac{9}{4}(u\,du^2 - 2du\,dv + v\,dv^2) \tag{1}$$

als gegeben vorausgesetzt. Da die sechs Größen $\xi, \eta, \zeta, \xi', \eta', \zeta'$ Integrale der gleichen Moutardschen Differentialgleichung:

$$\xi_{\alpha\beta} = e^{\theta}\,\xi \quad \text{bzw.} \quad \xi'_{\alpha\beta} = e^{\theta}\,\xi' \tag{2}$$

sind, so liegt es nahe, sie einer gemeinsamen Moutardschen Transformation zu unterwerfen, die durch eine weitere, noch näher zu bestimmende Lösung R der Differentialgleichung (2):

$$R_{\alpha\beta} = e^{\theta} R \tag{3}$$

definiert sein mag. Wir versuchen dann, über R so zu verfügen, daß sich die sechs transformierten Größen zur Konstruktion einer neuen Biegungsfläche verwenden lassen. Das gelingt nun, wie sich zeigen wird, insofern nicht unmittelbar, als man die beiden durch die Moutardsche Transformation gewonnenen Tripel, die wir mit $\bar\xi_1, \bar\eta_1, \bar\zeta_1$ und $\bar\xi_1', \bar\eta_1', \bar\zeta_1'$ bezeichnen wollen, erst mit je einem konstanten Faktor zu multiplizieren hat, damit sie in die den Forderungen von § 2 vollständig genügenden Größen ξ_1, η_1, ζ_1 und $\xi_1', \eta_1', \zeta_1'$ übergehen.

Wir schreiben zunächst die bekannten Differentialrelationen der Moutardschen Transformation[23])

[23]) Bianchi, Lezioni di geometria diff. 2 (1903), § 241.

$$(4)\quad \begin{cases} R(\xi_1)_\alpha + R_\alpha \xi_1 = -R\xi_\alpha + R_\alpha \xi, & R(\xi_1)_\beta + R_\beta \xi_1 = R\xi_\beta - R_\beta \xi, \\ R(\xi_1')_\alpha + R_\alpha \xi_1' = -R\xi_\alpha' + R_\alpha \xi', & R(\xi_1')_\beta + R_\beta \xi_1' = R\xi_\beta' - R_\beta \xi' \end{cases}$$

und verlangen, daß analog zu § 2 (3) wieder die Gleichungen:

$$(5)\quad \begin{cases} \sum \xi_1 \xi_1' = -1, & \sum (\xi_1)_\alpha \xi_1' = 0, & \sum (\xi_1)_\beta \xi_1' = 0, \\ & \sum (\xi_1')_\alpha \xi_1 = 0, & \sum (\xi_1')_\beta \xi_1 = 0 \end{cases}$$

erfüllt sind, die zum Ausdruck bringen, daß auch die transformierten Hilfsflächen (ξ_1, η_1, ζ_1) und $(\xi_1', \eta_1', \zeta_1')$ polarreziprok in bezug auf die imaginäre Einheitskugel um den Nullpunkt sind. Multipliziert man die beiden ersten Formeln (4) mit ξ_1', die beiden anderen mit ξ_1 und addiert jedesmal die analogen Beziehungen, so findet man mit Benutzung von (5):

$$(6)\quad \begin{cases} -R_\alpha = -R\sum \xi_1' \xi_\alpha + R_\alpha \sum \xi_1' \xi, & -R_\beta = R\sum \xi_1' \xi_\beta - R_\beta \sum \xi_1' \xi, \\ -R_\alpha = -R\sum \xi_1 \xi_\alpha' + R_\alpha \sum \xi_1 \xi', & -R_\beta = R\sum \xi_1 \xi_\beta' - R_\beta \sum \xi_1 \xi'. \end{cases}$$

Da $\Delta = (\xi, \xi_\alpha, \xi_\beta) = -e^\theta \neq 0$ ist, ist für ξ_1, η_1, ζ_1 ein Ansatz von der Form

$$\xi_1 = l\,\xi_\alpha + m\,\xi_\beta + n\,\xi \quad \text{usw.}$$

statthaft; entsprechend kann man

$$\xi_1 = l'\,\xi_\alpha' + m'\,\xi_\beta' + n'\,\xi' \quad \text{usw.}$$

setzen. Mittels der Relationen von § 2, 2 gewinnt man dann aus (6):

$$l = -(n+1)e^{-\theta}\frac{R_\beta}{R}, \qquad m = -(n-1)e^{-\theta}\frac{R_\alpha}{R},$$

$$l' = -(n'+1)e^{-\theta}\frac{R_\beta}{R}, \qquad m' = -(n'-1)e^{-\theta}\frac{R_\alpha}{R}.$$

Die erste der Relationen (5) läßt sich jetzt in die folgende verwandeln:

$$(nn'-1)\left(2e^{-\theta}\frac{R_\alpha R_\beta}{R^2} - 1\right) = 0.$$

Das identische Verschwinden der zweiten Klammer wäre mit der Differentialgleichung

$$(7)\qquad \theta_{\alpha\beta} = e^\theta - e^{-2\theta}$$

unvereinbar[24]); es muß also $n' = \frac{1}{n}$ sein. Demnach bestehen Formeln von der folgenden Gestalt:

---

[24]) Differentiiert man nämlich die Gleichung $2R_\alpha R_\beta = e^\theta R^2$ nach α, so ergibt sich: $2R_{\alpha\alpha} R_\beta = e^\theta \theta_\alpha R^2$, also: $R_{\alpha\alpha} = \theta_\alpha R_\alpha$. Differentiiert man abermals nach β, so wird

$$(e^\theta R)_\alpha = \theta_{\alpha\beta} R_\alpha + \theta_\alpha e^\theta R,$$

also: $\theta_{\alpha\beta} = e^\theta$ im Widerspruch mit (7).

$$(8)\qquad \begin{cases} \xi_1 = n\left\{\xi - \frac{e^{-\theta}}{R}\left[\left(1+\frac{1}{n}\right)R_\beta\xi_\alpha + \left(1-\frac{1}{n}\right)R_\alpha\xi_\beta\right]\right\}, \\ \xi_1' = \frac{1}{n}\left\{\xi' - \frac{e^{-\theta}}{R}\left[(1+n)R_\beta\xi_\alpha' + (1-n)R_\alpha\xi_\beta'\right]\right\}. \end{cases}$$

Setzt man den ersten dieser beiden Ausdrücke wieder in die Differentialrelationen (4) der Moutardschen Transformation ein, die man hierzu zweckmäßig in der Form

$$(R\xi_1)_\alpha + R\xi_\alpha - R_\alpha\xi = 0, \qquad (R\xi_1)_\beta - R\xi_\beta + R_\beta\xi = 0$$

schreibt, und drückt darauf die zweiten Ableitungen von ξ an der Hand von § 2 (10) durch ξ, ξ_α, ξ_β aus, so findet man, da wegen $\Delta \neq 0$ die Faktoren von ξ, ξ_α, ξ_β einzeln verschwinden müssen:

$$n_\alpha = 0, \qquad n_\beta = 0, \quad \text{also} \quad n = \text{konst.}$$

und außerdem zwei neue Differentialgleichungen für R, zu denen wir (3) wieder hinzunehmen:

$$(9)\qquad \begin{cases} R_{\alpha\alpha} = \theta_\alpha R_\alpha + \frac{1+n}{1-n}e^{-\theta}R_\beta, \\ R_{\alpha\beta} = e^\theta R, \\ R_{\beta\beta} = \frac{1-n}{1+n}e^{-\theta}R_\alpha + \theta_\beta R_\beta. \end{cases}$$

Das zweite Formelpaar (4) führt zu demselben Ergebnis. Die *willkürliche* Konstante n muß von 0 und 1 verschieden sein.

2. Es sei also R ein Integral des nach § 4, 2 unbeschränkt integrablen Systems (9), auf das Tzitzéica die Transformation der durch $\frac{K}{W^4} = \text{konst.}$ charakterisierten Flächen zurückgeführt hat. Es fragt sich jetzt, wie weit die nach (8) mit Hilfe von R gebildeten Größen $\xi_1, \ldots, \xi_1', \ldots$ den Bedingungen von § 2 entsprechen. Zunächst wissen wir, daß die Differentialrelationen (4) der Moutardschen Transformation *erfüllt* sind, und schließen auf Grund der bekannten Eigenschaften derselben, daß

$$(10)\qquad (\xi_1)_{\alpha\beta} = e^{\theta_1}\xi_1, \qquad (\xi_1')_{\alpha\beta} = e^{\theta_1}\xi_1'$$

wird, wobei

$$e^{\theta_1} = R\left(\frac{1}{R}\right)_{\alpha\beta} = -e^\theta + 2\frac{R_\alpha R_\beta}{R^2}$$

ist. Wir bestätigen mittels (8) die Beziehung

$$(11)\qquad \sum \xi_1\xi_1' = -1$$

und bilden die Hilfsformeln:

$$(12)\qquad \sum \xi_1\xi' = -n, \qquad \sum \xi_1'\xi = -\frac{1}{n},$$

$$(13)\quad \begin{cases} \sum \xi_1 \xi'_\alpha = -(n-1)\frac{R_\alpha}{R}, & \sum \xi_1 \xi'_\beta = -(n+1)\frac{R_\beta}{R}, \\ \sum \xi'_1 \xi_\alpha = -\left(\frac{1}{n}-1\right)\frac{R_\alpha}{R}, & \sum \xi'_1 \xi_\beta = -\left(\frac{1}{n}+1\right)\frac{R_\beta}{R}, \end{cases}$$

Wir multiplizieren nun das erste Gleichungspaar (4) mit ξ'_1, das zweite mit ξ_1, addieren jedesmal die beiden analogen Relationen und finden, indem wir (11), (12) und (13) beachten:

$$(14)\quad \Sigma(\xi_1)_\alpha \xi'_1 = 0, \quad \Sigma(\xi_1)_\beta \xi'_1 = 0, \quad \Sigma(\xi'_1)_\alpha \xi_1 = 0, \quad \Sigma(\xi'_1)_\beta \xi_1 = 0.$$

Damit sind die Formeln (5) verifiziert, die uns zur Spezialisiernng der Moutardschen Transformation gedient haben. Wir multiplizieren ferner jede Gleichung des esten Paares (4) mit jeder des zweiten, addieren wieder die analogen Relationen und erhalten nach entsprechenden Reduktionen:

$$(15)\quad \sum(\xi_1)_\alpha(\xi'_1)_\alpha = 0, \qquad \sum(\xi_1)_\beta(\xi'_1)_\beta = 0,$$

$$(16)\quad \sum(\xi_1)_\alpha(\xi'_1)_\beta = \sum(\xi_1)_\beta(\xi'_1)_\alpha = -e^\theta + 2\frac{R_\alpha R_\beta}{R^2} = e^{\theta_1}.$$

Auf den beiden neuen Flächen (ξ_1, η_1, ζ_1) und $(\xi'_1, \eta'_1, \zeta'_1)$ wird das Kurvennetz (α, β), wie aus (15) hervorgeht, wiederum von den Asymptotenlinien gebildet. Es bestehen demnach wie in § 2 Beziehungen von der Form:

$$(17)\quad (\xi_1)_{\alpha\alpha} = (\theta_1)_\alpha(\xi_1)_\alpha + p_1(\xi_1)_\beta,$$

und wir wissen, daß p_1 sich von $e^{-\theta_1}$ nur durch einen von α allein abhängigen Faktor unterscheiden kann. Es ist von Wichtigkeit, festzustellen, daß $p_1 = e^{-\theta_1}$ ist. Zu diesem Zwecke differentiieren wir die erste der Differentialrelationen (4) nach α und drücken dabei $(\xi_1)_{\alpha\alpha}$ durch (17) sowie $\xi_{\alpha\alpha}$ durch § 2 (10) aus. Multiplizieren wir sodann mit $(\xi'_1)_\alpha$ und addieren die beiden analogen Gleichungen, so folgt mit Berücksichtigung von (14), (15) und (16):

$$(18)\quad R p_1 e^{\theta_1} = -R\theta_\alpha \Sigma(\xi'_1)_\alpha \xi_\alpha - R e^{-\theta}\Sigma(\xi'_1)_\alpha \xi_\beta + R_{\alpha\alpha}\Sigma(\xi'_1)_\alpha \xi.$$

Die drei hier auftretenden Summen erhält man, indem man für $(\xi'_1)_\alpha$ den durch (4) gegebenen Ausdruck benutzt und von den Formeln (13) Gebrauch macht:

$$\sum(\xi'_1)_\alpha \xi_\alpha = \frac{1-n}{n}\left(\frac{R_\alpha}{R}\right)^2, \qquad \sum(\xi'_1)_\alpha \xi_\beta = \frac{1+n}{n}\frac{R_\alpha R_\beta}{R^2} - e^\theta.$$

$$\sum(\xi'_1)_\alpha \xi = -\sum \xi'_1 \xi_\alpha = \frac{1-n}{n}\frac{R_\alpha}{R}.$$

Setzt man diese Werte, ebenso den aus (9) entnommenen Ausdruck für $R_{\alpha\alpha}$ in (18) ein, so ergibt sich tatsächlich $p_1 = e^{-\theta_1}$.

Auf die gleiche Weise bilden wir $(\bar{\xi}_1)_{\beta\beta}$ und erhalten so das System:

$$(\xi_1)_{\alpha\alpha} = (\theta_1)_\alpha (\bar{\xi}_1)_\alpha + e^{-\theta_1}(\bar{\xi}_1)_\beta, \qquad (\bar{\xi}_1)_{\alpha\beta} = e^{\theta_1}\bar{\xi}_1,$$
$$(\bar{\xi}_1)_{\beta\beta} = e^{-\theta_1}(\xi_1)_\alpha + (\theta_1)_\beta(\bar{\xi}_1)_\beta,$$

aus dem wir unter Bezugnahme auf § 2 unmittelbar schließen, daß

$$(\theta_1)_{\alpha\beta} = e^{\theta_1} - e^{-2\theta_1}$$

wird, daß also *durch*

(19) $$e^{\theta_1} = -e^{\theta} + 2\frac{R_\alpha R_\beta}{R^2}$$

ein neues Integral θ_1 *der partiellen Differentialgleichung* (7) *gegeben ist*[25]).

Um endlich den Wert der Determinante

$$\bar{\Delta}_1 = (\bar{\xi}_1, (\bar{\xi}_1)_\alpha, (\bar{\xi}_1)_\beta)$$

zu berechnen, bestimmen wir $(\bar{\xi}_1)_\alpha$ und $(\bar{\xi}_1)_\beta$ aus (4) und ersetzen nachträglich $\bar{\xi}_1$ durch den Ausdruck (8). Wir erhalten:

$$\bar{\Delta}_1 = \left(n\xi - (n+1)e^{-\theta}\frac{R_\beta}{R}\xi_\alpha - (n-1)e^{-\theta}\frac{R_\alpha}{R}\xi_\beta,\ \frac{R_\alpha}{R}\xi - \xi_\alpha,\ -\frac{R_\beta}{R}\xi + \xi_\beta\right)$$

$$= (\xi, \xi_\alpha, \xi_\beta)\begin{vmatrix} n, & -(n+1)e^{-\theta}\frac{R_\beta}{R}, & -(n-1)e^{-\theta}\frac{R_\alpha}{R} \\ \frac{R_\alpha}{R}, & -1, & 0 \\ \frac{R_\beta}{R}, & 0, & 1 \end{vmatrix} = -n\Delta\left(2e^{-\theta}\frac{R_\alpha R_\beta}{R^2} - 1\right)$$

und weiter, indem wir (19) sowie die Relation $\Delta = -e^\theta$ beachten:

(20) $$\bar{\Delta} = -ne^{\theta_1}.$$

Analog wird:

(21) $$\bar{\Delta}_1' = (\bar{\xi}_1', (\bar{\xi}_1')_\alpha, (\bar{\xi}_1')_\beta) = -\frac{1}{n}e^{\theta_1},$$

wovon wir uns am einfachsten durch Multiplikation der beiden Determinanten $\bar{\Delta}_1$ und $\bar{\Delta}_1'$ überzeugen.

Die Formeln (20) und (21) unterscheiden sich von § 2 (11) und (15) durch das Auftreten der Faktoren n und $\frac{1}{n}$. Wir lassen deshalb an die

---

[25]) Dieser Satz wurde bereits von Tzitzéica [s. die erste der unter [1]) zitierten Noten] ausgesprochen. Den direkten rechnerischen Nachweis findet man in § 3 meiner unter [8]) genannten Arbeit.

Stelle der durch die gemeinsame Moutardsche Transformation gewonnenen $\bar{\xi}_1, \ldots, \bar{\xi}_1', \ldots$ die neuen Größen:

$$
(22) \qquad \begin{cases} \xi_1 = n^{-1/2}\bar{\xi}_1, & \eta_1 = n^{-1/2}\bar{\eta}_1, & \zeta_1 = n^{-1/2}\bar{\zeta}_1, \\ \xi_1' = n^{1/2}\bar{\xi}_1', & \eta_1' = n^{1/2}\bar{\eta}_1', & \zeta_1 = n^{1/2}\bar{\zeta}_1' \end{cases}
$$

treten, für die die auf entsprechende Weise gebildeten Determinanten $\Delta_1 = \Delta_1' = -e^{\theta_1}$ werden. *Überhaupt gelten jetzt die Relationen von* § 2 *ausnahmslos für den Index* 1.

Hiernach sind die Koordinaten der neuen Biegungsfläche (x_1, y_1, z_1) vom Typus (1), wenn

$$\Sigma \xi_1^2 = u_1, \quad \Sigma \xi_1'^2 = v_1$$

gesetzt wird, zunächst durch die Formeln

$$(23) \qquad x_1 = \frac{3}{2}\int(\xi_1 du_1 + \xi_1' dv_1) \quad \text{usw.}$$

gegeben. In bezug auf (22) ist zu bemerken, daß nach § 4, **1** kein unmittelbarer Anlaß vorzuliegen scheint, die drei Größen eines jeden der beiden Tripel mit dem gleichen normierenden Faktor $n^{-1/2}$ bzw. $n^{1/2}$ zu versehen. Tatsächlich erweist sich aber gerade dieser Schritt als bedeutungsvoll. Er gestattet nämlich, wie jetzt gezeigt werden soll, die Beseitigung der Quadraturen (23), die eine wesentliche Voraussetzung für ein tieferes Eindringen in die geometrische Natur unserer Transformation darstellt.

3. Wir führen unter Hinweis auf § 1, **3** und § 3, **3** die zu der neuen Biegungsfläche (x_1, y_1, z_1) gehörige *dritte Hilfsfläche* $(\mathfrak{x}_1, \mathfrak{y}_1, \mathfrak{z}_1)$ ein, die durch

$$(24) \qquad \mathfrak{x}_1 = x_1 - u_1\xi_1 - v_1\xi_1'$$

gegeben ist und nach § 3 (5) die Darstellung:

$$\mathfrak{x}_1 = \langle \eta_1, \zeta_1' \rangle + \langle \eta_1', \zeta_1 \rangle$$

zuläßt. Es sei zunächst festgestellt, daß die letzte Formel sich auch in der Gestalt:

$$(25) \qquad \mathfrak{x}_1 = \langle \bar{\eta}_1, \bar{\zeta}_1' \rangle + \langle \bar{\eta}_1', \bar{\zeta}_1 \rangle$$

schreiben läßt. Erinnert sei überdies an diejenige fundamentale Eigenschaft der Moutardschen Transformation, auf die Guichard die Konstruktion der W-Kongruenzen zu einem gegebenen Brennflächenmantel gegründet hat. Sind nämlich φ und ψ zwei Integrale einer Moutardschen Gleichung und φ_1, ψ_1 die Ergebnisse einer gemeinsamen Moutardschen Transformation[26]),

[26]) Ist R die transformierende Funktion, so können die Ergebnisse der Moutardschen Transformation mittels unseres Symbols in der Form $\varphi_1 = \frac{1}{R}\langle R, \varphi \rangle$, $\psi_1 = \frac{1}{R}\langle R, \psi \rangle$ geschrieben werden.

so kann in dem Ausdruck $\langle \varphi_1, \psi_1 \rangle$ über die additive Integrationskonstante so verfügt werden, daß

$$\langle \varphi_1, \psi_1 \rangle = \langle \varphi, \psi \rangle + \varphi_1 \psi - \psi_1 \varphi \tag{26}$$

wird [27]).

Danach folgt aus (25) im Verein mit § 3 (5):

$$\begin{aligned} \mathfrak{x}_1 &= \mathfrak{x} + \bar{\eta}_1 \zeta' - \bar{\zeta}_1 \eta' + \bar{\eta}'_1 \zeta - \bar{\zeta}'_1 \eta \\ &= \mathfrak{x} + n^{1/3}(\eta_1 \zeta' - \zeta_1 \eta') + n^{-1/3}(\eta'_1 \zeta - \zeta'_1 \eta), \end{aligned} \tag{27}$$

so daß man x_1 mittels (24) ohne Quadraten erhält.

Wir sind jetzt in der Lage, den Übergang von der gegebenen zur neuen Biegungsfläche, den wir als *Transformation* Θ_n bezeichnen, zusammenfassend, wie folgt, zu beschreiben:

Es sei (x, y, z) *eine auf ihre Asymptotenlinien* (α, β) *bezogene Biegungsfläche vom Typus*

$$ds^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2);$$

es seien $\xi = \frac{2}{3} x_u, \ldots, \xi' = \frac{2}{3} x_v, \ldots$ *die* 6 *Hilfsgrößen, die sich als Koordinaten zweier polarreziproker Flächen* (ξ, η, ζ) *und* (ξ', η', ζ') *der Tzitzéicaschen Klasse auffassen lassen. Ist* R *ein Integral des unbeschränkt integrablen Systems*:

$$R_{\alpha\alpha} = \theta_\alpha R_\alpha + \frac{1+n}{1-n} e^{-\theta} R_\beta, \quad R_{\alpha\beta} = e^{\theta} R, \quad R_{\beta\beta} = \frac{1-n}{1+n} e^{-\theta} R_\alpha + \theta_\beta R_\beta, \tag{28}$$

in dem n *eine von* 0 *und* 1 *verschiedene Konstante bedeutet, so bilde man* ξ_1, η_1, ζ_1 *und* $\xi'_1, \eta'_1, \zeta'_1$ *auf Grund der Formeln*:

$$\left\{ \begin{aligned} \xi_1 &= n^{2/3} \left\{ \xi - \frac{e^{-\theta}}{R} \left[\left(1 + \frac{1}{n}\right) R_\beta \xi_\alpha + \left(1 - \frac{1}{n}\right) R_\alpha \xi_\beta \right] \right\}, \text{[28]} \\ \xi'_1 &= n^{-2/3} \left\{ \xi' - \frac{e^{-\theta}}{R} \left[(1+n) R_\beta \xi'_\alpha + (1-n) R_\alpha \xi'_\beta \right] \right\}. \end{aligned} \right. \tag{29}$$

---

[27]) Aus drei Lösungen der Moutardschen Gleichung und ihren Transformierten erhält man die beiden auf die Asymptotenlinien bezogenen Flächen $x = \langle \eta, \zeta \rangle, \ldots$ und $x_1 = \langle \eta_1, \zeta_1 \rangle, \ldots$. Aus (26) folgt dann: $x_1 - x = \eta_1 \zeta - \zeta_1 \eta$ usw., d. h. die Tatsache, daß (x, y, z) und (x_1, y_1, z_1) die Brennflächenmäntel einer Kongruenz sind.

[28]) Wir haben bereits erwähnt, daß nach Tzitzéica [s. wieder unter [1])] von den Differentialgleichungen 28) die *asymptotische* Transformation abhängt, die eine Fläche (ξ, η, ζ) der durch die Relation $\dfrac{K}{W^4} = \text{konst.}$ ausgezeichneten Klasse in eine Fläche der gleichen Art — sie sei hier mit $(\hat{\xi}_1, \hat{\eta}_1, \hat{\zeta}_1)$ bezeichnet — überführt. Für die Berechnung der transformierten Fläche gibt Tzitzéica die Formel:

$$\hat{\xi}_1 = \xi - \frac{e^{-\theta}}{R} \left[\left(1 + \frac{1}{n}\right) R_\beta \xi_\alpha + \left(1 - \frac{1}{n}\right) R_\alpha \xi_\beta \right]$$

(Fortsetzung der Fußnote 28 auf nächster Seite)

Wird

$$\sum \xi_1^2 = u_1, \quad \sum \xi_1'^2 = v_1$$

gesetzt, so stellen

$$(3) \quad \begin{cases} x_1 = x + u_1\xi_1 + v_1\xi_1' - u\xi - v\xi' + n^{1/2}(\eta_1\zeta' - \zeta_1\eta') + n^{-1/2}(\eta_1'\zeta - \zeta_1'\eta), \\ y_1 = y + u_1\eta_1 + v_1\eta_1' - u\eta - v\eta' + n^{1/2}(\zeta_1\xi' - \xi_1\zeta') + n^{-1/2}(\zeta_1'\xi - \xi_1'\zeta), \\ z_1 = z + u_1\zeta_1 + v_1\zeta_1' - u\zeta - v\zeta' + n^{1/2}(\xi_1\eta' - \eta_1\xi') + n^{-1/2}(\xi_1'\eta - \eta_1'\xi) \end{cases}$$

die laufenden Koordinaten einer neuen Biegungsfläche vom gleichen Typus dar, für die

$$dx_1 = \frac{3}{2}(\xi_1\, du_1 + \xi_1'\, dv_1) \quad \text{usw.}$$

und

$$\sum dx_1^2 = \frac{9}{4}(u_1\, du_1^2 - 2\, du_1\, dv_1 + v_1\, dv_1^2)$$

wird.

Was den Grad der Allgemeinheit der Transformation Θ_n anbetrifft, so ist zu beachten, daß das System (28) drei linear-unabhängige Integrale besitzt, mittels derer sich jedes weitere Integral in der Form

$$R = c_1 R_1 + c_2 R_2 + c_3 R_3$$

darstellen läßt. Die Formeln (29) enthalten nur die Verhältnisse der drei Konstanten. Bei *gegebenem* n liefert also die Transformation Θ_n zu einer gegebenen Biegungsfläche ∞^2 neue. Ohne nähere Ausführungen sei noch die an der Hand der voraufgehenden Entwicklungen leicht erweisbare Tatsache ausgesprochen, daß *die zu* Θ_n *inverse Transformation eine* $\Theta_{\frac{1}{n}}$ *ist.* Insbesondere genügt also $\frac{1}{R}$ einem System von Differentialgleichungen, das aus (28) hervorgeht, wenn θ durch θ_1 und n durch $\frac{1}{n}$ ersetzt wird.

4. Die naheliegende, aber auf direktem Wege durchaus nicht mühelos zu erledigende Frage, ob es sich bei der Transformation Θ_n vielleicht auch wieder um eine asymptotische Transformation handelt, braucht uns hier nicht zu beschäftigen, da die Ergebnisse des Schlußparagraphen genügen, um sie im negativen Sinne zu beantworten. Eine erste, höchst interessante geometrische Eigenschaft der Transformation Θ_n knüpft sich an die Ein-

---

$\left(\text{es steht dort } c \text{ an Stelle von } \frac{1}{n}\right)$, so daß also $\hat{\xi}_1 = n^{-3/2}\xi_1 = \frac{1}{n}\bar{\xi}_1$ ist. Es ist unmittelbar zu ersehen, daß der Punkt $(\hat{\xi}_1, \hat{\eta}_1, \hat{\zeta}_1)$ auf einer Tangente der Fläche (ξ, η, ζ) liegt; daß auch die umgekehrte Beziehung statthat, daß also die Flächen $(\xi, \ldots)$ und $(\hat{\xi}_1, \ldots)$ Brennflächen einer Strahlenkongruenz sind, erkennt man leicht, indem man mittels (13) bestätigt, daß $\sum(\hat{\xi}_1 - \xi)\,\bar{\xi}_1' = 0$ wird. Hervorgehoben sei übrigens, daß der Wert des Bruches $\frac{K}{W^4}$, den wir für die gegebene Fläche $(\xi, \ldots)$ gleich -1 voraussetzen konnten, für die Tzitzéicasche Transformierte notwendig ein anderer, nämlich $-n^4$ wird.

führung der Flächen $(x^{(\nu)}, \ldots)$, deren Punkte, wie erinnerlich, mit den Flächenelementen der Biegungsfläche $(x, \ldots)$ starr gekoppelt, in den Tangentialebenen liegen. Es gilt nun der folgende Satz:

Eine jede mit der gegebenen Biegungsfläche $(x, \ldots)$ *verbundene Fläche* $(x^{(\nu)}, \ldots)$ *und eine zu der transformierten Biegungsfläche* $(x_1, \ldots)$ *gehörige Fläche* $(x_1^{(\nu_1)}, \ldots)$, *für die*

$$\nu_1 = \nu n^{-2/4}$$

ist, bilden die Brennflächen einer W-Kongruenz. Wir können auch sagen: *Bei der Transformation* Θ_n *gehen die sämtlichen Flächen* $(x^{(\nu)}, \ldots)$ *durch simultane asymptotische Transformationen in die Flächen* $(x_1^{(\nu_1)}, \ldots)$ *über.*

Wir beachten nämlich, daß nach § 3 (6) und (7):

$$(31)\quad \begin{cases} x^{(\nu)} = x - (u - \nu)\,\xi - \left(v - \dfrac{1}{\nu}\right)\xi' = \mathfrak{x} + \nu\xi + \dfrac{1}{\nu}\xi' \\ \qquad = \dfrac{1}{\nu}\langle \eta + \nu\eta', \zeta + \nu\zeta' \rangle, \\ x_1^{(\nu_1)} = x_1 - (u_1 - \nu_1)\,\xi_2 - \left(v_1 - \dfrac{1}{\nu_1}\right)\xi_1' = \mathfrak{x}_1 + \nu_1\xi_1 + \dfrac{1}{\nu_1}\xi_1' \\ \qquad = \dfrac{1}{\nu_1}\langle \eta_1 + \nu_1\eta_1', \zeta_1 + \nu_1\zeta_1' \rangle \end{cases}$$

ist. Der Beweis beruht nun auf der unmittelbar zu übersehenden Tatsache, daß $\xi + \nu\xi'$ durch die zu R gehörige Moutardsche Transformation in $\bar{\xi}_1 + \nu\bar{\xi}_1' = n^{1/4}\xi_1 + \nu n^{-1/4}\xi_1'$, also in $n^{1/4}(\xi_1 + \nu_1\xi_1')$ übergeführt wird, wenn $\nu n^{-2/4} = \nu_1$ gesetzt wird. Man hat sich aber noch davon zu überzeugen, daß die relative Lage der Flächen $(x^{(\nu)}, \ldots)$ und $(x_1^{(\nu_1)}, \ldots)$ wirklich eine derartige ist, daß die Verbindungslinien ihrer korrespondierenden Punkte gemeinschaftliche Tangenten sind. Hierzu bedarf es der auch später zu benutzenden Relationen:

$$(32)\quad n^{-1/4}\xi_1 - n^{1/4}\xi = \eta_1'\zeta' - \zeta_1'\eta', \qquad n^{1/4}\xi_1' - n^{-1/4}\xi' = \eta_1\zeta - \zeta_1\eta,$$

die man unschwer als Identitäten feststellt, wenn man die den Index 1 tragenden Größen mittels der Formeln (29) ausdrückt. Man folgert nun aus (31):

$$x_1^{(\nu_1)} - x^{(\nu)} = \mathfrak{x}_1 - \mathfrak{x} + \nu(n^{-2/4}\xi_1 - \xi) + \frac{1}{\nu}(n^{2/4}\xi_1' - \xi')$$

und erhält mit Benutzung von (27) und (32):

$$\begin{aligned} x_1^{(\nu_1)} - x^{(\nu)} = n^{1/4}(\eta_1\zeta' - \zeta_1\eta') + n^{-1/4}(\eta_1'\zeta - \zeta_1'\eta) + \nu n^{-1/4}(\eta_1'\zeta' - \zeta_1'\eta') \\ + \frac{1}{\nu}n^{1/4}(\eta_1\zeta - \zeta_1\eta), \end{aligned}$$

also:

$$x_1^{(\nu_1)} - x^{(\nu)} = \frac{n^{1/4}}{\nu}\left[(\eta_1 + \nu_1\eta_1')(\zeta + \nu\zeta') - (\zeta_1 + \nu_1\zeta_1')(\eta + \nu\eta')\right]$$

woraus die Richtigkeit der Behauptung hervorgeht.

§ 6.

Der Vertauschbarkeitssatz für die Transformation Θ_n.

1. Zur Aufstellung des Vertauschbarkeitssatzes für unsere Transformation Θ_n gehen wir von dem allgemeinen Bianchischen *Kompositionstheorem für die Moutardschen Transformationen* aus[29]). Dasselbe wird zweckmäßig in der folgenden Form ausgesprochen:

Es seien $\varphi, \psi, \ldots$ beliebig viele Integrale einer Moutardschen Differentialgleichung. Durch zwei weitere Lösungen R_1 und R_2 seien zwei Moutardsche Transformationen definiert, durch die die gegebenen Größen in $\varphi_1, \psi_1, \ldots$ und $\varphi_2, \psi_2, \ldots$ übergeführt werden. Bestimmt man J mittels des Ausdrucks

$$(1)\quad J=\langle R_1, R_2\rangle=\int\{-[R_1(R_2)_\alpha-R_2(R_1)_\alpha]\,d\alpha+[R_1(R_2)_\beta-R_2(R_1)_\beta]\,d\beta]\},$$

der eine willkürliche additive Konstante enthält, so gehören zu den Größen

$$(2)\qquad R_{13}=\frac{J}{R_1},\quad R_{23}=\frac{J}{R_2},$$

die Lösungen der von $\varphi_1, \psi_1, \ldots$ bzw. von $\varphi_2, \psi_2, \ldots$ erfüllten Moutardschen Gleichungen sind, wiederum zwei Moutardsche Transformationen, durch die aus $\varphi_1, \psi_1, \ldots$ einerseits, aus $\varphi_2, \psi_2, \ldots$, andrerseits einunddieselbe Reihe $\varphi_3, \psi_3, \ldots$ von Integralen einer vierten Moutardschen Gleichung hervorgeht. Dabei berechnen sich $\varphi_3, \psi_3, \ldots$ nach der Formel:

$$(3)\qquad \varphi_3=\varphi-\frac{R_1R_2}{J}(\varphi_1-\varphi_2).$$

Überdies gilt der folgende Zusatz: *Bei geeigneter Fixierung der dem Symbol $\langle\ \rangle$ anhaftenden Integrationskonstanten bestehen neben den Relationen*

$$(4)\quad \langle\varphi_1, \psi_1\rangle=\langle\varphi, \psi\rangle+\varphi_1\psi-\psi_1\varphi,\quad \langle\varphi_2, \psi_2\rangle=\langle\varphi, \psi\rangle+\varphi_2\psi-\psi_2\varphi$$

die folgenden:

$$(5)\quad \langle\varphi_3, \psi_3\rangle=\langle\varphi_1, \psi_1\rangle+\varphi_3\psi_1-\psi_3\varphi_1=\langle\varphi_2, \psi_2\rangle+\varphi_3\psi_2-\psi_3\varphi_2,$$

an deren Stelle wegen (3) *und* (4) *die Gleichung*

$$(6)\qquad \langle\varphi_3, \psi_3\rangle=\langle\varphi, \psi\rangle-\frac{R_1R_2}{J}(\varphi_1\psi_2-\psi_1\varphi_2)$$

treten kann.

Auf diesen Zusatz gründet sich die geometrische Deutung des Kompositionstheorems im Falle dreier gegebener Integrale ξ, η, ζ. Sie führt bekanntlich zur Konstruktion der Quadrupel von W-Kongruenzen, die aus ∞^2 windschiefen Vierecken bestehen, deren Ecken die einander zugeordneten

[29]) Bianchi, Lezioni di geometria diff. 2 (1903), §§ 247 u. 248.

Punkte und deren Seiten Tangenten von vier Flächen mit korrespondierenden Asymptotenlinien bilden. Bei dem uns beschäftigenden Problem handelt es sich um die Transformationen eines Systems von sechs Integralen $\xi, \eta, \zeta, \xi', \eta', \zeta'$. Auch hier ist der leitende Gedanke wieder der, statt mit den Biegungsflächen selber zunächst mit den zugehörigen *dritten* Hilfsflächen zu operieren, für die die erweiterten Lelieuvreschen Formeln § 3 (5) gelten.

2. Die betrachtete Biegungsfläche $(x, \ldots)$ vom Typus § 5 (1) werde also zwei Transformationen Θ_{n_1} und Θ_{n_2} unterworfen, die durch zwei Integrale R_1 und R_2 des für $n = n_1$ und $n = n_2$ geschriebenen Systems (28) von § 5 definiert sein mögen. Dann ist nach § 5 (29):

$$(7)\quad \begin{cases} \xi_1 = n_1^{1/2} \left\{\xi - \dfrac{e^{-\theta}}{R_1}\left[\left(1+\dfrac{1}{n_1}\right)(R_1)_\beta\,\xi_\alpha + \left(1-\dfrac{1}{n_1}\right)(R_1)_\alpha\,\xi_\beta\right]\right\}, \\ \xi_1' = n_1^{-1/2}\left\{\xi' - \dfrac{e^{-\theta}}{R_1}\left[(1+n_1)(R_1)_\beta\,\xi_\alpha' + (1-n_1)(R_1)_\alpha\,\xi_\beta'\right]\right\}, \\ \xi_2 = n_2^{1/2} \left\{\xi - \dfrac{e^{-\theta}}{R_2}\left[\left(1+\dfrac{1}{n_2}\right)(R_2)_\beta\,\xi_\alpha + \left(1-\dfrac{1}{n_2}\right)(R_2)_\alpha\,\xi_\beta\right]\right\}, \\ \xi_2' = n_2^{-1/2}\left\{\xi' - \dfrac{e^{-\theta}}{R_2}\left[(1+n_2)(R_2)_\beta\,\xi_\alpha' + (1-n_2)(R_2)_\alpha\,\xi_\beta'\right]\right\}; \end{cases}$$

dazu kommen die Formeln:

$$\mathfrak{x}_1 = \mathfrak{x} + n_1^{1/2}(\eta_1\zeta' - \zeta_1\eta') + n_1^{-1/2}(\eta_1'\zeta - \zeta_1'\eta),$$
$$x_1 = \mathfrak{x}_1 + u_1\,\xi_1 + v_1\,\xi_1' \qquad (u_1 = \Sigma\xi_1^2, \quad v_1 = \Sigma\xi_1'^2)$$

und entsprechende für den Index 2.

Es soll nun gezeigt werden, daß *bei geeigneter Festsetzung der in J enthaltenen verfügbaren Konstante* die Anwendung des allgemeinen Kompositionstheorems für die Moutardschen Transformationen sechs Größen $\xi_3, \eta_3, \zeta_3, \xi_3', \eta_3', \zeta_3'$ liefert, aus denen sich eine vierte Biegungsfläche $(x_3, \ldots)$ konstruieren läßt, und außerdem, daß diese vierte Biegungsfläche $(x_3, \ldots)$ aus $(x_1, \ldots)$ mittels einer Transformation Θ_{n_2} und aus $(x_2, \ldots)$ mittels einer Θ_{n_1} hervorgeht. Die charakteristischen Konstanten n_1 und n_2 erscheinen dabei also *vertauscht*.

Auf die normierenden Faktoren ist besonders Bedacht zu nehmen. Es geht nämlich durch die mit den ursprünglichen Transformationen Θ_{n_1} und Θ_{n_2} gleichbedeutenden, zu R_1 und R_2 gehörigen Moutardschen Transformationen

$$\xi \text{ in } n_1^{1/2}\xi_1 \text{ und in } n_2^{1/2}\xi_2,$$
$$\xi' \text{ in } n_1^{-1/2}\xi_1' \text{ und in } n_2^{-1/2}\xi_2'$$

über. Das veranlaßt uns, von vornherein das gemeinsame Ergebnis der beiden weiteren, von R_{13} und R_{23} (siehe (2)) abhängigen Moutardschen

Transformationen, denen $n_1^{1/2}\xi_1$ und $n_1^{-1/2}\xi_1'$ bzw. $n_2^{1/2}\xi_2$ und $n_2^{-1/2}\xi_2'$ unterworfen werden, mit $n_1^{1/2}n_2^{1/2}\xi_3$ und $n_1^{-1/2}n_2^{-1/2}\xi_3'$ zu bezeichnen. Wir definieren also an Hand der Formeln (1), (2) und (3) ξ_3 und ξ_3' durch die Ausdrücke:

$$
(8)\qquad \left\{\begin{aligned}
\xi_3 &= n_1^{-1/2}n_2^{-1/2}\left[\xi - \frac{R_1 R_2}{J}\left(n_1^{1/2}\xi_1 - n_2^{1/2}\xi_2'\right)\right],\\
\xi_3' &= n_1^{1/2}n_2^{1/2}\left[\xi' - \frac{R_1 R_2}{J}\left(n_1^{-1/2}\xi_1' - n_2^{-1/2}\xi_2'\right)\right].
\end{aligned}\right.
$$

Der Beweis des Vertauschbarkeitssatzes wird erbracht sein, wenn es uns gelingt, diese Größen als Ergebnisse zweier Transformationen Θ_{n_1} und Θ_{n_2}, also in jeder der beiden folgenden Formen darzustellen:

$$
(9)\qquad \left\{\begin{aligned}
\xi_3 &= n_2^{3/2}\left\{\xi_1 - \frac{e^{-\theta_1}}{R_{13}}\left[\left(1+\frac{1}{n_2}\right)(R_{13})_\beta(\xi_1)_\alpha + \left(1-\frac{1}{n_2}\right)(R_{13})_\alpha(\xi_1)_\beta\right]\right\},\\
\xi_3' &= n_2^{-3/2}\left\{\xi_1' - \frac{e^{-\theta_1}}{R_{13}}\left[(1+n_2)(R_{13})_\beta(\xi_1')_\alpha + (1-n_2)(R_{13})_\alpha(\xi_1')_\beta\right]\right\},\\
\xi_3 &= n_1^{3/2}\left\{\xi_2 - \frac{e^{-\theta_2}}{R_{23}}\left[\left(1+\frac{1}{n_1}\right)(R_{23})_\beta(\xi_2)_\alpha + \left(1-\frac{1}{n_1}\right)(R_{13})_\alpha(\xi_2)_\beta\right]\right\},\\
\xi_3' &= n_1^{-3/2}\left\{\xi_2' - \frac{e^{-\theta_2}}{R_{23}}\left[(1+n_1)(R_{23})_\beta(\xi_2')_\alpha + (1-n_1)(R_{23})_\alpha(\xi_2')_\beta\right]\right\}.
\end{aligned}\right.
$$

Es genügt dann nämlich, die zu § 5 (4) analogen, infolge des allgemeinen Kompositionstheorems sicher bestehenden Differentialrelationen der durch R_{13} und R_{23} definierten Moutardschen Transformationen hinzuzunehmen, um im Hinblick auf die Entwicklungen von § 5 unmittelbar schließen zu können, daß R_{13} und R_{23} die beiden Systeme von Differentialgleichungen:

$$
(R_{13})_{\alpha\alpha} = (\theta_1)_\alpha(R_{13})_\alpha + \frac{1+n_2}{1-n_2}e^{-\theta_1}(R_{13})_\beta \quad \text{usw.},
$$
$$
(R_{23})_{\alpha\alpha} = (\theta_2)_\alpha(R_{23})_\alpha + \frac{1+n_1}{1-n_1}e^{-\theta_1}(R_{23})_\beta \quad \text{usw.}
$$

erfüllen, durch die die beiden Transformationen als eine Θ_{n_2} und eine Θ_{n_1} gekennzeichnet werden.

Wir begnügen uns damit, von den fraglichen Formeln (9) die erste auf Grund der dem J noch aufzuerlegenden Bedingung zu bestätigen; bei den übrigen hätte man entsprechend zu verfahren. Berücksichtigt man, daß $\Delta_1' = (\xi_1', (\xi_1')_\alpha, (\xi_1')_\beta) = -e^{\theta_1} \neq 0$ ist und daß die zu § 2 (3) und (4) analogen Relationen

$$
\sum \xi_1\xi_1' = -1, \qquad \sum \xi_1(\xi_1')_\alpha = \sum \xi_1(\xi_1')_\beta = 0,
$$
$$
\sum (\xi_1)_\alpha(\xi_1')_\alpha = \sum (\xi_1)_\beta(\xi_1')_\beta = 0, \qquad \sum (\xi_1)_\alpha(\xi_1')_\beta = \sum (\xi_1)_\beta(\xi_1')_\alpha = e^{\theta_1}
$$

bestehen, so wird ersichtlich, daß man die zu bestätigende Formel für ξ_3

samt den beiden entsprechenden für η_3 und ζ_3 durch das vollkommen gleichwertige System der drei Gleichungen:

$$
(10)\quad \begin{cases} \sum \xi_3 \xi_1' = -n_2^{2/3}, \\ \sum \xi_3 (\xi_1')_\alpha = -n_2^{2/3}\left(1-\frac{1}{n_2}\right)\frac{(R_{13})_\alpha}{R_{13}}, \quad \sum \xi_3 (\xi_1')_\beta = -n_2^{2/3}\left(1+\frac{1}{n_2}\right)\frac{(R_{13})_\beta}{R_{13}} \end{cases}
$$

ersetzen kann. Es handelt sich also darum, durch Einführen des Ausdrucks (8) für ξ_3 hieraus Identitäten zu machen. Um die erforderlichen Reduktionen vornehmen zu können, gebrauchen wir eine Reihe von Hilfsformeln. Man findet zunächst:

$$
(11)\quad \sum \xi \xi_1 = -n_1^{-2/3}, \quad \sum \xi_1 \xi_1' = -1, \quad \sum \xi_2 \xi_1' = -n_1^{1/3} n_2^{-1/3} + n_1^{-2/3} n_2^{-1/3} \frac{\Psi}{R_1 R_2},
$$

wobei

$$
(12)\quad \begin{aligned} \Psi = {} & (1-n_1 n_2) e^{-\theta} [(R_1)_\alpha (R_2)_\beta - (R_1)_\beta (R_2)_\alpha] \\ & + (n_1 - n_2)[R_1 R_2 - e^{-\theta}((R_1)_\alpha (R_2)_\beta + (R_1)_\beta (R_2)_\alpha)] \end{aligned}
$$

ist. Es wird ferner:

$$
(13)\quad \begin{cases} \sum \xi (\xi_1')_\alpha = -\sum \xi_\alpha \xi_1' = n_1^{-2/3}(1-n_1)\frac{(R_1)_\alpha}{R_1}, \\ \sum \xi (\xi_1')_\beta = n_1^{-2/3}(1+n_1)\frac{(R_1)_\beta}{R_1}. \end{cases}
$$

Wir bilden schließlich $\sum \xi_2 (\xi_1')_\alpha$ und $\sum \xi_2 (\xi_1')_\beta$, indem wir berücksichtigen, daß nach § 5 (4) und (22)

$$
(14)\quad \begin{cases} (\xi_1')_\alpha = -\frac{(R_1)_\alpha}{R_1}\xi_1' - n_1^{1/3}\left(\xi_\alpha' - \frac{(R_1)_\alpha}{R_1}\xi'\right), \\ (\xi_1')_\beta = -\frac{(R_1)_\beta}{R_1}\xi_1' + n_1^{1/3}\left(\xi_\beta' - \frac{(R_1)_\beta}{R_1}\xi'\right) \end{cases}
$$

ist, und uns alsdann zur Ausführung der Summen der letzten Formel (11) sowie der folgenden leicht zu bestätigenden Beziehungen bedienen:

$$
\sum \xi_2 \xi' = -n_2^{2/3}, \quad \sum \xi_2 \xi_\alpha' = -n_2^{2/3}\left(1-\frac{1}{n_2}\right)\frac{(R_2)_\alpha}{R_2},
$$

$$
\sum \xi_2 \xi_\beta' = -n_2^{2/3}\left(1+\frac{1}{n_2}\right)\frac{(R_2)_\beta}{R_2}.
$$

Es ergibt sich so das Formelpaar

$$
(15)\quad \begin{cases} \sum \xi_2 (\xi_1')_\alpha = -n_1^{1/3} n_2^{2/3}\left(1-\frac{1}{n_2}\right)\frac{J_\alpha}{R_1 R_2} - n_1^{-2/3} n_2^{-1/3}\frac{(R_1)_\alpha}{R_1^2 R_2}\Psi, \\ \sum \xi_2 (\xi_1')_\beta = -n_1^{1/3} n_2^{2/3}\left(1+\frac{1}{n_2}\right)\frac{J_\beta}{R_1 R_2} - n_1^{-2/3} n_2^{-1/3}\frac{(R_1)_\beta}{R_1^2 R_2}\Psi, \end{cases}
$$

wobei im Hinblick auf (1):

$$
(16)\quad -[R_1 (R_2)_\alpha - R_2 (R_1)_\alpha] = J_\alpha, \quad R_1 (R_2)_\beta - R_2 (R_1)_\beta = J_\beta
$$

gesetzt wurde. Nach diesen Vorbereitungen substituieren wir den Wert (8) von ξ_3 in die drei Summen, die die linken Seiten der fraglichen Relationen (10) bilden. Wird dann von den Hilfsformeln (11) und (15) Gebrauch gemacht und berücksichtigt, daß nach (2):

$$\frac{J}{R_1} = R_{13}, \qquad \frac{J}{R_2} = R_{23}$$

ist, so gelingt es, die Summen (10) in die rechts stehenden Ausdrücke überzuführen, wofern die Bedingung

$$\Psi = (1 - n_1 n_2) J \tag{17}$$

erfüllt ist. Damit wird aber, wenn wir zunächst die Annahme $n_1 n_2 = 1$ ausschließen, lediglich über die dem J anhaftende additive Konstante verfügt, da, wie wir durch Differentiation von (12) bestätigen, die Beziehung

$$d\Psi = (1 - n_1 n_2) dJ$$

von selber besteht. An die Stelle der Quadratur (2) tritt also zur Berechnung von J der endliche Ausdruck:

$$J = \frac{\Psi}{1 - n_1 n_2}.$$

3. Um nun die Formeln für die laufenden Koordinaten der durch $\xi_3, \ldots, \xi_3', \ldots$ bestimmten Biegungsfläche (x_3, y_3, z_3) zu erhalten, gehen wir wieder von der zugehörigen *dritten* Hilfsfläche $(\mathfrak{x}_3, \ldots)$ aus, die, wenn $u_3 = \sum \xi_3^2$ und $v_3 = \sum \xi_3'^2$ ist, sich in der Form

$$\mathfrak{x}_3 = x_3 - u_3 \xi_3 - v \xi_3' = \langle \eta_3, \zeta_3' \rangle + \langle \eta_3', \zeta_3 \rangle$$

darstellen läßt. Auf Grund von (5) und (6) finden wir:

$$\begin{aligned} \mathfrak{x}_3 &= \mathfrak{x}_1 + n_2^{1/2}(\eta_3 \zeta_1' - \zeta_3 \eta_1') + n_2^{-1/2}(\eta_3' \zeta_1 - \zeta_3' \eta_1) \\ &= \mathfrak{x}_2 + n_1^{1/2}(\eta_3 \zeta_2' - \zeta_3 \eta_2') + n_1^{-1/2}(\eta_3' \zeta_2 - \zeta_3' \eta_2) \\ &= \mathfrak{x} - \frac{R_1 R_2}{J}\left[n_1^{1/2} n_2^{-1/2}(\eta_1 \zeta_2' - \zeta_1 \eta_2') + n_1^{-1/2} n_2^{1/2}(\eta_1' \zeta_2 - \zeta_1' \eta_2)\right]. \end{aligned}$$

Wir können demnach den *Vertauschbarkeitssatz für die Transformation* Θ_n in der folgenden Gestalt aussprechen:

Wird eine Biegungsfläche (x, y, z) *vom Typus*

$$ds^2 = \frac{9}{4}(u\,du^2 - 2\,du\,dv + v\,dv^2)$$

mittels zweier durch R_1 *bzw.* R_2 *definierter Transformationen* Θ_{n_1} *und* Θ_{n_2} *in die beiden neuen Biegungsflächen* (x_1, y_1, z_1) *und* (x_2, y_2, z_2) *übergeführt, so existiert* ($n_1 n_2 \neq 1$ *vorausgesetzt*) *eine vierte, durch endliche Operationen bestimmbare Biegungsfläche* (x_3, y_3, z_3), *die aus* (x_1, y_1, z_1)

durch eine Transformation Θ_{n_2} und aus (x_2, y_2, z_2) durch eine Θ_{n_1} hervorgeht. Dabei ist:

$$(18)\qquad J = e^{-\theta}[(R_1)_\alpha (R_2)_\beta - (R_1)_\beta (R_2)_\alpha]$$
$$+ \frac{n_1 - n_2}{1 - n_1 n_2}[R_1 R_2 - e^{-\theta}((R_1)_\alpha (R_2)_\beta + (R_1)_\beta (R_2)_\alpha)],$$
$$\xi_3 = n_1^{-1/2} n_2^{-1/2}\left[\xi - \frac{R_1 R_2}{J}(n_1^{1/2}\xi_1 - n_2^{1/2}\xi_2)\right],$$
$$\xi_3' = n_1^{1/2} n_2^{1/2}\left[\xi' - \frac{R_1 R_2}{J}(n_1^{-1/2}\xi_1' - n_2^{-1/2}\xi_2')\right],\quad u_3 = \Sigma \xi_3^2,\quad v_3 = \Sigma \xi_3'^2,$$
$$x_3 = x + u_3 \xi_3 + v_3 \xi_3' - u\xi - v\xi'$$
$$- \frac{R_1 R_2}{J}[n_1^{1/2} n_2^{-1/2}(\eta_1 \zeta_2' - \zeta_1 \eta_2') + n_1^{-1/2} n_2^{1/2}(\eta_1' \zeta_2 - \zeta_1' \eta_2)].$$

Im Falle $n_1 = n_2$ reduziert sich (18) auf den einfacheren Ausdruck:

$$J = e^{-\theta}[(R_1)_\alpha (R_2)_\beta - (R_1)_\beta (R_2)_\alpha].$$

Es ist also auch dann noch eine einzige, durch endliche Operationen bestimmbare vierte Biegungsfläche $(x_3, \ldots)$ vorhanden.

Ist dagegen $n_1 n_2 = 1$, so versagt die Formel (18). Wir ersehen aber aus der ursprünglichen Form (17) der Relation, daß dann $\Psi = 0$, also

$$(19)\qquad R_1 R_2 - e^{-\theta}[(R_1)_\alpha (R_2)_\beta + (R_1)_\beta (R_2)_\alpha] = 0$$

sein muß. Der auf der linken Seite stehende Ausdruck ist aber, wie man durch Differentiieren oder an Hand einer in § 4, 2 gemachten Bemerkung erkennt, für $n_1 = n$, $n_2 = \frac{1}{n}$ eine Konstante. Die Relation (19) hat somit die Bedeutung einer Anfangsbedingung, die bei der Integration des einen der beiden von R_1 oder von R_2 zu erfüllenden Systeme von Differentialgleichungen zu berücksichtigen ist. Zwei durch die Bedingung (19) miteinander verknüpfte Transformationen Θ_n und $\Theta_{\frac{1}{n}}$ wollen wir, indem wir uns eines bei ähnlichen Verhältnissen gelegentlich benutzten Ausdruckes bedienen, als ein *harmonisches* Paar bezeichnen.

Es besteht hiernach der folgende *Zusatz* zum Vertauschbarkeitssatz:

In dem besonderen Falle $n_1 n_2 = 1$ $\left(n_1 = n,\ n_2 = \frac{1}{n}\right)$ gilt der Vertauschbarkeitssatz für die harmonischen, d. h. durch das Bestehen der Relation (19) *ausgezeichneten Paare von Transformationen Θ_n* und $\Theta_{\frac{1}{n}}$ *und verlangt hier die Ausführung der Quadratur*

$$J = \int \{-[R_1 (R_2)_\alpha - R_2 (R_1)_\alpha]\, d\alpha + [R_1 (R_2)_\beta - R_2 (R_1)_\beta]\, d\beta\}.$$

Da die Integrationskonstante willkürlich bleibt, gibt es eine einfach unendliche Schar von vierten Biegungsflächen (x_3, y_3, z_3).

Auf der Betrachtung solcher harmonischer Paare von Transformationen beruhen die Entwicklungen des folgenden Schlußparagraphen, die uns bemerkenswerte Aufschlüsse über die geometrische Natur der Transformation Θ_n geben werden.

§ 7.

Die Existenz dreigliedriger Transformationszyklen und die Zerlegung der Transformation Θ_n in zwei sukzessive asymptotische Transformationen.

1. Wir setzen wieder eine Biegungsfläche (x, y, z) der betrachteten Klasse als gegeben voraus und unterwerfen sie zwei durch R_1 und R_2 definierten Transformationen Θ_n und $\Theta_{\frac{1}{n}}$, die ein harmonisches Paar bilden mögen. Es ist dann also:

$$(1)\qquad R_1 R_2 - e^{-\theta}[(R_1)_\alpha (R_2)_\beta + (R_1)_\beta (R_2)_\alpha] = 0$$

und:

$$(2)\qquad \begin{cases} \xi_1 = n^{1/2} \left\{\xi - \dfrac{e^{-\theta}}{R_1}\left[\left(1+\dfrac{1}{n}\right)(R_1)_\beta \xi_\alpha + \left(1-\dfrac{1}{n}\right)(R_1)_\alpha \xi_\beta\right]\right\}, \\ \xi_1' = n^{-1/2}\left\{\xi' - \dfrac{e^{-\theta}}{R_1}[(1+n)(R_1)_\beta \xi_\alpha' + (1-n)(R_1)_\alpha \xi_\beta']\right\}, \\ \xi_2 = n^{-1/2}\left\{\xi - \dfrac{e^{-\theta}}{R_2}[(1+n)(R_2)_\beta \xi_\alpha + (1-n)(R_2)_\alpha \xi_\beta]\right\}, \\ \xi_2' = n^{1/2} \left\{\xi' - \dfrac{e^{-\theta}}{R_2}\left[\left(1+\dfrac{1}{n}\right)(R_2)_\beta \xi_\alpha' + \left(1-\dfrac{1}{n}\right)(R_2)_\alpha \xi_\beta'\right]\right\}. \end{cases}$$

Die zugehörigen, nach § 5 (30) darstellbaren transformierten Biegungsflächen seien (x_1, y_1, z_1) und (x_2, y_2, z_2).

Es soll nun gezeigt werden, daß durch

$$(3)\qquad R_{12} = \frac{R_2}{R_1}$$

eine weitere Transformation Θ_n bestimmt ist, die $(x_1, \ldots)$ in $(x_2, \ldots)$ verwandelt, daß also die drei Biegungsflächen $(x, \ldots)$, $(x_1, \ldots)$, $(x_2, \ldots)$ durch einen dreigliedrigen Zyklus von Transformationen Θ_n miteinander verbunden sind. Zu diesem Zweck haben wir zunächst die Beziehungen:

$$(4)\qquad \begin{cases} \xi_2 = n^{1/2}\left\{\xi_1 - \dfrac{e^{-\theta_1}}{R_{12}}\left[\left(1+\dfrac{1}{n}\right)(R_{12})_\beta (\xi_1)_\alpha + \left(1-\dfrac{1}{n}\right)(R_{12})_\alpha (\xi_1)_\beta\right]\right\}, \\ \xi_2' = n^{-1/2}\left\{\xi_1' - \dfrac{e^{-\theta_1}}{R_{12}}\left[(1+n)(R_{12})_\beta (\xi_1')_\alpha + (1-n)(R_{12})_\alpha (\xi_1')_\beta\right]\right\} \end{cases}$$

zu bestätigen. Wir beschränken uns darauf, den Beweis für die erste Formel durchzuführen, die wir zusammen mit den entsprechenden für η_2

und ξ_2, indem wir ganz ähnlich wie im vorigen Paragraphen verfahren, durch das folgende gleichwertige System ersetzen:

$$\sum \xi_2 \xi_1' = -n^{1/2},$$

$$\sum \xi_2 (\xi_1')_\alpha = -n^{1/2}\left(1 - \frac{1}{n}\right)\frac{(R_{12})_\alpha}{R_{12}}, \quad \sum \xi_2 (\xi_1')_\beta = -n^{1/2}\left(1 + \frac{1}{n}\right)\frac{(R_{12})_\beta}{R_{12}}.$$

Die Richtigkeit dieser Relationen ergibt sich aber unmittelbar aus § 6 (11), (15) und (16), wenn berücksichtigt wird, daß jetzt $n_1 = n$, $n_2 = \frac{1}{n}$ und $\Psi = 0$ ist.

An zweiter Stelle überzeugen wir uns davon, daß ξ_1 und $n^{1/2}\xi_2$ sowie ξ_1' und $n^{-1/2}\xi_2'$ die Differentialrelationen einer durch $R_{12} = \frac{R_2}{R_1}$ definierten Moutardschen Transformation erfüllen. Wir schreiben von den fraglichen Formeln nur die erste hin:

$$(5) \qquad (\xi_2)_\alpha + \frac{(R_{12})_\alpha}{R_{12}}\xi_2 = -n^{-1/2}\left[(\xi_1)_\alpha - \frac{(R_{12})_\alpha}{R_{12}}\xi_1\right].$$

Benutzt man, um hierin $(\xi_1)_\alpha$ und $(\xi_2)_\alpha$ auszudrücken, die analogen Beziehungen:

$$(\xi_1)_\alpha + \frac{(R_1)_\alpha}{R_1}\xi_1 = -n^{-1/2}\left[\xi_\alpha - \frac{(R_1)_\alpha}{R_1}\xi\right],$$

$$(\xi_2)_\alpha + \frac{(R_2)_\alpha}{R_2}\xi_2 = -n^{1/2}\left[\xi_\alpha - \frac{(R_2)_\alpha}{R_2}\xi\right],$$

die infolge der gegebenen Transformationen Θ_n und $\Theta_{\frac{1}{n}}$ bestehen, und ersetzt schließlich ξ_1 und ξ_2 durch die Werte (2), so erweist sich (5) als Identität. Entsprechendes gilt für die übrigen Formeln der Gruppe.

Unter Bezugnahme auf die Entwicklungen von § 5, **1** schließen wir aus dem gleichzeitigen Bestehen der Formeln (4) und der Differentialrelationen der Moutardschen Transformation, daß eine zwischen den Biegungsflächen $(x_1, \ldots)$ und $(x_2, \ldots)$ vermittelnde Transformation Θ_n vorliegt, insbesondere also, daß $R_{12} = \frac{R_2}{R_1}$ das System der Differentialgleichungen (28) von § 5 befriedigt, in dem θ_1 an Stelle von θ zu schreiben ist.

Es ist indessen noch der Nachweis zu erbringen, daß die Biegungsfläche $(x_2, \ldots)$, die aus $(x_1, \ldots)$ durch die neue Transformation Θ_n hervorgeht, auch ihrer Lage nach mit der Fläche $(x_2, \ldots)$ zusammenfällt, die wir ursprünglich aus $(x, \ldots)$ durch die Transformation $\Theta_{\frac{1}{n}}$ erhielten. Die Integraldarstellung der Koordinaten:

$$x_2 = \frac{3}{2}\int (\xi_2\, du_2 + \xi_2'\, dv_2) \qquad \left(u_2 = \sum \xi_2^2, \quad v_2 = \sum \xi_2'^2\right)$$

schließt nämlich noch die Möglichkeit einer Translation ein. Wir haben

also, indem wir uns wieder der mit den Biegungsflächen verbundenen *dritten* Hilfsflächen bedienen, zu zeigen, daß neben den Relationen:

$$(6)\quad \begin{cases} \mathfrak{x}_1 - \mathfrak{x} = n^{1/2}(\eta_1\zeta' - \zeta_1\eta') + n^{-1/2}(\eta_1'\zeta - \zeta_1'\eta), \\ \mathfrak{x} - \mathfrak{x}_2 = n^{1/2}(\eta\zeta_2' - \zeta\eta_2') + n^{-1/2}(\eta'\zeta_2 - \zeta'\eta_2) \end{cases}$$

auch die folgende besteht:

$$(7)\quad \mathfrak{x}_2 - \mathfrak{x}_1 = n^{1/2}(\eta_2\zeta_1' - \zeta_2\eta_1') + n^{-1/2}(\eta_2'\zeta_1 - \zeta_2'\eta_1).$$

Es verschwindet aber tatsächlich beim Addieren dieser drei Gleichungen die rechts entstehende Summe, da einzeln die beiden Identitäten:

$$(8)\quad \begin{cases} n^{1/2}(\eta\zeta_2' + \eta_1\zeta' + \eta_2\zeta_1') - n^{-1/2}(\eta\zeta_1' + \eta_1\zeta_2' + \eta_2\zeta') = 0, \\ n^{1/2}(\zeta\eta_2' + \zeta_1\eta' + \zeta_2\eta_1') - n^{-1/2}(\zeta\eta_1' + \zeta_1\eta_2' + \zeta_2\eta') = 0 \end{cases}$$

erfüllt sind, von denen man z. B. die erste leicht bestätigt, indem man ihre linke Seite in der Form:

$$\eta(n^{1/2}\zeta_2' - n^{-1/2}\zeta_1') + \eta_1(n^{1/2}\zeta' - n^{-1/2}\zeta_2') + \eta_2(n^{1/2}\zeta_1' - n^{-1/2}\zeta') = 0$$

schreibt und die nach § 5 (32) geltenden Beziehungen:

$$(9)\quad \begin{cases} n^{1/2}\zeta_1' - n^{-1/2}\zeta' = \xi_1\eta - \eta_1\xi, \quad n^{1/2}\zeta' - n^{-1/2}\zeta_2' = \xi\eta_2 - \eta\xi_2, \\ \qquad n^{1/2}\zeta_2' - n^{-1/2}\zeta_1' = \xi_2\eta_1 - \eta_2\xi_1 \end{cases}$$

benutzt.

Wir stellen das hiermit gewonnene Ergebnis fest: *Gehen aus einer Biegungsfläche* (x, y, z) *des betrachteten Typus durch ein harmonisches Paar von Transformationen* Θ_n *und* $\Theta_{\frac{1}{n}}$ *die beiden neuen Biegungsflächen* (x_1, y_1, z_1) *und* (x_2, y_2, z_2) *hervor, so vermittelt auch zwischen* (x_1, y_1, z_1) *und* (x_2, y_2, z_2) *eine Transformation* Θ_n. *Es liegt demnach ein dreigliedriger Zyklus von Transformationen* Θ_n, *im umgekehrten Sinne von Transformationen* $\Theta_{\frac{1}{n}}$ *vor. Das Produkt der die drei Transformationen* Θ_n *definierenden Funktionen* R_1, $R_{12} = \frac{R_2}{R_1}$ *und* $\frac{1}{R_2}$ hat den *Wert* 1.

2. An die Konstruktion eines solchen dreigliedrigen Zyklus, dem eine beliebige gegebene Transformation Θ_n angehört, knüpfen wir nun den Beweis der höchst beachtenswerten Tatsache, daß die Transformation Θ_n sich in zwei aufeinander folgende asymptotische Transformationen auflösen läßt. Wir beginnen diesen letzten Teil unserer Untersuchung, bei dem längere, im folgenden zum Teil nur angedeutete Rechnungen unvermeidlich erscheinen, mit der Ermittlung *des Schnittpunkts* (x^*, y^*, z^*) *der korrespondierenden Tangentialebenen* der drei zuletzt betrachteten

Biegungsflächen $(x, \ldots), (x_1, \ldots)$ und $(x_2, \ldots)$. Man findet auf Grund der Relationen (6) und (7):

$$\begin{aligned}\mathfrak{x}_1 - \mathfrak{x} &= -[(\mathfrak{x}_2 - \mathfrak{x}_1) + (\mathfrak{x} - \mathfrak{x}_2)] \\ &= -[\eta_2(n^{1/2}\zeta_1' - n^{-1/2}\zeta') + \eta_2'(n^{-1/2}\zeta_1 - n^{1/2}\zeta) \\ &\qquad - \zeta_2(n^{1/2}\eta_1' - n^{-1/2}\eta') - \zeta_2'(n^{-1/2}\eta_1 - n^{1/2}\eta)]\end{aligned}$$

und, indem man sich der Gleichungen (9) sowie weiterer analog gebildeter Formeln bedient, nach leicht zu übersehender Umformung:

$$\mathfrak{x}_1 - \mathfrak{x} = \xi \sum \xi_1 \xi_2 + \xi' \sum \xi_1' \xi_2' - \xi_1 \sum \xi \xi_2 - \xi_1' \sum \xi' \xi_2'.$$

Demnach besteht Gleichheit der beiden folgenden Ausdrücke, deren gemeinsamen Wert wir mit x^* bezeichnen:

$$\mathfrak{x} + \xi \sum \xi_1 \xi_2 + \xi' \sum \xi_1' \xi_2' = \mathfrak{x}_1 + \xi_1 \sum \xi \xi_2 + \xi_1' \sum \xi' \xi_2' = x^*.$$

Man erkennt unmittelbar, daß gleichzeitig auch:

$$\mathfrak{x}_2 + \xi_2 \sum \xi \xi_1 + \xi_2' \sum \xi' \xi_1' = x^*$$

ist. Führen wir die Koordinaten der Biegungsflächen ein, indem wir uns der Beziehung § 1 (15) erinnern, so ergibt sich:

$$(10) \qquad \begin{cases} x^* = x + (\sum \xi_1 \xi_2 - u)\xi + (\sum \xi_1' \xi_2' - v)\xi' \\ \quad\;\; = x_1 + (\sum \xi \xi_2 - u_1)\xi_1 + (\sum \xi' \xi_2' - v)\xi_1' \\ \quad\;\; = x_2 + (\sum \xi \xi_1 - u_2)\xi_2 + (\sum \xi \xi_1' - v_2)\xi_2'. \end{cases}$$

Diese Formeln, denen analoge für y^* und z^* hinzuzufügen sind, lassen erkennen, daß *der Punkt* (x^*, y^*, z^*) *gleichzeitig in korrespondierenden Tangentialebenen der drei Biegungsflächen liegt.*

3. Es wäre nun zu untersuchen, ob die Verbindungslinien des Punktes $(x^*, \ldots)$ mit den entsprechenden Punkten der drei Biegungsflächen auch Tangenten der von $(x^*, \ldots)$ beschriebenen Fläche sind. Will man die Frage auf dem nächstliegenden Wege, der in der Bestimmung der Flächennormalen von $(x^*, \ldots)$ bestehen würde, entscheiden, so stößt man auf erhebliche rechnerische Komplikationen, deren Ursache in den drei verschiedenen Darstellungen (10) von x^* liegt. Es sei deshalb zunächst diejenige Tatsache festgestellt, aus der ich auf die Zerlegbarkeit der Transformation Θ_n schließen mußte. Sie gab überdies die Richtlinien für die auf die Moutardsche Transformation zu gründende Entwicklung.

Aus einer Bemerkung in § 4, 2 entnehmen wir, daß die von $\Xi = \eta\zeta' - \zeta\eta', \ldots$ erfüllte Moutardsche Gleichung:

$$(11) \qquad \Xi_{\alpha\beta} = 3 e^{\theta} \Xi$$

als partikuläre Lösung das Produkt $R_1 R_2$ zuläßt, wenn R_1 und R_2 ein

durch die Relation (1) ausgezeichnetes, also harmonisches Paar von Transformationen Θ_n und $\Theta_{\frac{1}{n}}$ definieren. Es ist danach möglich, Ξ, H, Z durch die zu $R_1 R_2$ gehörige Moutardsche Transformation in die Lösungen Ξ^*, H*, Z* einer neuen Moutardschen Gleichung:

$$\Xi^*_{\alpha\beta} = \mathsf{M}^* \Xi^*$$

überzuführen. Dabei wird auf Grund einer bekannten Beziehung:

$$\mathsf{M}^* = -3e^\theta + 2\frac{(R_1 R_2)_\alpha (R_1 R_2)_\beta}{R_1^2 R_2^2}.$$

Führt man die Differentiationen aus und berücksichtigt dabei die Gleichungen:

$$e^{\theta_1} + e^\theta = 2\frac{(R_1)_\alpha (R_1)_\beta}{R_1^2}, \qquad e^{\theta_2} + e^\theta = 2\frac{(R_2)_\alpha (R_2)_\beta}{R_2^2}$$

sowie die harmonische Relation (1), so findet man:

$$\mathsf{M}^* = e^\theta + e^{\theta_1} + e^{\theta_2}.$$

Die symmetrische Gestalt dieses Ausdrucks legt die Vermutung nahe, daß die drei Biegungsflächen $(x, \ldots)$, $(x_1, \ldots)$ und $(x_2, \ldots)$ eine gemeinsame asymptotische Transformierte zulassen. Der Nachweis dieses Sachverhalts wird durch den Umstand erschwert, daß die Ergebnisse

$$\Xi^* = \frac{1}{R_1 R_2}\langle R_1 R_2, \Xi\rangle \quad \text{usw.} \tag{12}$$

der durch $R_1 R_2$ definierten Moutardschen Transformation Quadraturen enthalten und infolgedessen auch nicht eindeutig sind.

4. Wir bilden die zu Ξ, H, Z analogen Größen:

$$\Xi_1 = \eta_1 \zeta_1' - \zeta_1 \eta_1', \ldots, \qquad \Xi_2 = \eta_2 \zeta_2' - \zeta_2 \eta_2', \ldots,$$

die den Normalenkosinus der Biegungsflächen $(x_1, \ldots)$ und $(x_2, \ldots)$ proportional sind. Trifft nun die soeben ausgesprochene Vermutung zu, so müssen Ξ^*, H*, Z*, für die an Stelle von (12) endliche und eindeutige Ausdrücke zu suchen sind, ebenso wie aus Ξ, H, Z auch aus Ξ_1, H$_1$, Z$_1$ und aus Ξ_2, H$_2$, Z$_2$ durch Moutardsche Transformationen hervorgehen, für die die charakteristischen Funktionen, dem Produkte $R_1 R_2$ entsprechend, bezüglich $R_{12}\frac{1}{R_1} = \frac{R_2}{R_1^2}$ und $\frac{1}{R_{12}}\frac{1}{R_2} = \frac{R_1}{R_2^2}$ sind. Man hat sich also davon zu überzeugen, daß die Differentialrelationen:

$$(13)\quad \begin{cases} \Xi_\alpha^* = -\,\Xi_\alpha \quad - \left[\dfrac{(R_1)_\alpha}{R_1} + \dfrac{(R_2)_\alpha}{R_2}\right](\Xi^* - \Xi) \\ \quad = -\,(\Xi_1)_\alpha - \left[\dfrac{(R_2)_\alpha}{R_2} - 2\,\dfrac{(R_1)_\alpha}{R_1}\right](\Xi^* - \Xi_1) \\ \quad = -\,(\Xi_2)_\alpha - \left[\dfrac{(R_1)_\alpha}{R_1} - 2\,\dfrac{(R_2)_\alpha}{R_2}\right](\Xi^* - \Xi_2) \end{cases}$$

und ebenso die Beziehungen zwischen den Ableitungen nach β:

$$(14)\quad \Xi_\beta^* = \Xi_\beta - \left[\frac{(R_1)_\beta}{R_1} + \frac{(R_2)_\beta}{R_2}\right](\Xi^* + \Xi) \quad \text{usw.}$$

sich gleichzeitig erfüllen lassen. Um zunächst aus (13) einen *brauchbaren* Ausdruck für Ξ^* zu gewinnen, addieren wir die drei Gleichungen und erhalten:

$$(15)\quad \begin{cases} 3\,\Xi_\alpha^* = -\,(\Xi + \Xi_1 + \Xi_2)_\alpha + \left[\dfrac{(R_1)_\alpha}{R_1} + \dfrac{(R_2)_\alpha}{R_2}\right]\Xi \\ \qquad + \left[\dfrac{(R_2)_\alpha}{R_2} - 2\,\dfrac{(R_1)_\alpha}{R_1}\right]\Xi_1 + \left[\dfrac{(R_1)_\alpha}{R_1} - 2\,\dfrac{(R_2)_\alpha}{R_2}\right]\Xi_2. \end{cases}$$

Es ist andererseits:

$$(16)\quad n^{-1/2}(\eta\zeta_1' - \zeta\eta_1')_\alpha = -\,(\eta\zeta_\alpha' - \zeta\eta_\alpha') - [(\eta_1)_\alpha\zeta_1' - (\zeta_1)_\alpha\eta_1'] + \frac{(R_1)_\alpha}{R_1}(\Xi - \Xi_1).$$

Man beweist diese Beziehung an Hand der Formelgruppen

$$(17)\quad \begin{cases} (\xi_1)_\alpha + \dfrac{(R_1)_\alpha}{R_1}\,\xi_1 = -\,n^{-1/2}\left[\xi_\alpha - \dfrac{(R_1)_\alpha}{R_1}\,\xi\right] \quad \text{usw.,} \\ (\xi_1')_\alpha + \dfrac{(R_1)_\alpha}{R_1}\,\xi_1' = -\,n^{1/2}\left[\xi_\alpha' - \dfrac{(R_1)_\alpha}{R_1}\,\xi'\right] \quad \text{usw.,} \end{cases}$$

indem man links die Differentiation ausführt und η_α, ζ_α mittels $(\eta_1)_\alpha$, $(\zeta_1)_\alpha$ und umgekehrt $(\eta_1')_\alpha$, $(\zeta_1')_\alpha$ mittels η_α' und ζ_α' ausdrückt. Ebenso wird:

$$(18)\quad \begin{cases} n^{-1/2}(\eta_1\zeta_2' - \zeta_1\eta_2')_\alpha = -\,[\eta_1(\zeta_1')_\alpha - \zeta_1(\eta_1')_\alpha] - [(\eta_2)_\alpha\zeta_2' - (\zeta_2)_\alpha\eta_2'] \\ \qquad + \left[\dfrac{(R_2)_\alpha}{R_2} - \dfrac{(R_1)_\alpha}{R_1}\right](\Xi_1 - \Xi_2), \\ n^{-1/2}(\eta_2\zeta' - \zeta_2\eta')_\alpha = -\,[\eta_2(\zeta_2')_\alpha - \zeta_2(\eta_2')_\alpha] - (\eta_\alpha\zeta' - \zeta_\alpha\eta') \\ \qquad - \dfrac{(R_2)_\alpha}{R_2}(\Xi_2 - \Xi). \end{cases}$$

Addiert man die Gleichungen (16) und (18), so erkennt man, daß man die rechte Seite von (15) in der Form

$$n^{-1/3}(\eta\zeta_1' - \zeta\eta_1' + \eta_1\zeta_2' - \zeta_1\eta_2' + \eta_2\zeta' - \zeta_2\eta')_\alpha$$

schreiben kann. Wir schließen daraus, indem wir eine Integrationskonstante, deren Verschwinden kaum zweifelhaft ist, unterdrücken, daß

$$(19)\qquad \begin{aligned} \Xi^* &= \frac{n^{-1/3}}{3}(\eta\zeta_1' - \zeta\eta_1' + \eta_1\zeta_2' - \zeta_1\eta_2' + \eta_2\zeta' - \zeta_2\eta') \\ &= \frac{n^{1/3}}{3}(\eta\zeta_2' - \zeta\eta_2' + \eta_1\zeta' - \zeta_1\eta' + \eta_2\zeta_1' - \zeta_2\eta_1') \end{aligned}$$

ist. Die Identität des hinzugefügten zweiten Ausdrucks mit dem gefundenen ersten ist eine Folge der Gleichungen (8).

5. Es ist nun der Nachweis zu führen, daß die durch (19) definierte Größe Ξ^* den beiden Gruppen von Differentialrelationen (13) und (14) tatsächlich genügt. Wir beschränken uns darauf, den Gang der etwas mühseligen Rechnung für die erste der Gleichungen (13) anzudeuten. Die Symmetrie des Ausdrucks (19) bürgt dafür, daß dann auch die beiden anderen erfüllt sind. Bezüglich der zweiten, die Ableitungen nach β enthaltenden Formelgruppe (14) wäre entsprechend zu verfahren.

Eine nicht unerhebliche Abkürzung läßt sich erzielen, wenn man die Möglichkeit der Affintransformation von § 4, 1 berücksichtigt und erwäg daß der Ausdruck

$$(20)\qquad t = \frac{n^{-1/3}}{3}(\eta\zeta_1' + \eta_1\zeta_2' + \eta_2\zeta')$$

die mittels $R_1 R_2$ gewonnene Moutardsche Transformierte des Produkts $\eta\zeta'$ sein wird, von dem ja bekannt ist, daß es Integral von (11) ist. Es genügt hiernach offenbar, das Bestehen der Relation

$$(21)\qquad t_\alpha + \left[\frac{(R_1)_\alpha}{R_1} + \frac{(R_2)_\alpha}{R_2}\right] t = -(\eta\zeta')_\alpha + \left[\frac{(R_1)_\alpha}{R_1} + \frac{(R_2)_\alpha}{R_2}\right]\eta\zeta'$$

zu beweisen und nachträglich die Buchstaben η und ζ zu vertauschen. Wir beseitigen auf der linken Seite der fraglichen Gleichung (21), nachdem wir (20) differentiiert haben, mittels (17) sowie analoger Formeln für den Index 2 die sämtlichen Ableitungen der die Indizes 1 und 2 tragenden Größen. Zur weiteren Reduktion ist (1) heranzuziehen. Es ergibt sich zunächst:

$$(22)\qquad \begin{aligned} &t_\alpha + \left[\frac{(R_1)_\alpha}{R_1} + \frac{(R_2)_\alpha}{R_2}\right] t \\ &\qquad = \frac{1}{3}\left\{-(\eta\zeta')_\alpha + \left[\frac{(R_1)_\alpha}{R_1} + \frac{(R_2)_\alpha}{R_2}\right]\eta\zeta'\right\} + P + Q, \end{aligned}$$

wobei

$$P = \frac{n^{-1/3}}{3}[\eta_a(\zeta_1' - n^{-1/3}\zeta_2') + \zeta_a'(\eta_2 - n^{-1/3}\eta_1)],$$

$$Q = \frac{n^{-1/3}}{3}\left\{\eta\left[n^{-1/3}\frac{(R_1)_a}{R_1}\zeta_2' + \frac{(R_2)_a}{R_2}\zeta_1'\right] + \zeta'\left[\frac{(R_1)_a}{R_1}\eta_2 + n^{-1/3}\frac{(R_2)_a}{R_2}\eta_1\right]\right\}$$

ist. In P und Q haben wir die sämtlichen Größen mit dem Index 1 oder 2 durch die nach (2) gebildeten Ausdrücke zu ersetzen. Zu beachten ist außerdem § 2 (20). Man findet schließlich:

$$P = \frac{1}{3}\left(1 - \frac{1}{n}\right)\left\{-(\eta\zeta')_a + \left[\frac{(R_1)_a}{R_1} + \frac{(R_2)_a}{R_2}\right]\eta\zeta'\right\},$$

$$Q = \frac{1}{3}\left(1 + \frac{1}{n}\right)\left\{-(\eta\zeta')_a + \left[\frac{(R_1)_a}{R_1} + \frac{(R_2)_a}{R_2}\right]\eta\zeta'\right\}.$$

Werden diese Werte in (22) eingeführt, so bestätigt sich (21).

6. Ist nun wirklich eine gemeinsame asymptotische Transformierte (x^*, y^*, z^*) der drei Biegungsflächen vorhanden[30]), so muß diese mit dem ebenso bezeichneten, durch (10) dargestellten Ort der Schnittpunkte je dreier entsprechender Tangentialebenen zusammenfallen. Man übersieht leicht, indem man sich der bekannten Formeln für die Konstruktion der W-Kongruenzen erinnert[31]), daß es sich noch darum handelt, die Beziehung

$$x^* = x + (\sum \xi_1\xi_2 - u)\xi + (\sum \xi_1'\xi_2' - v)\xi' = x + 3(\mathsf{H}^*\mathsf{Z} - \mathsf{Z}^*\mathsf{H})$$

zu bestätigen, zu der man dann, ohne daß es eines besonderen Beweises hierfür bedarf, die beiden analogen Darstellungen von x^* hinzufügen kann:

$$x^* = x_1 + 3(\mathsf{H}^*\mathsf{Z}_1 - \mathsf{Z}^*\mathsf{H}_1) = x_2 + 3(\mathsf{H}^*\mathsf{Z}_2 - \mathsf{Z}^*\mathsf{H}_2).$$

Bilden wir H^* und Z^* nach (19), so finden wir:

$$3(\mathsf{H}^*\mathsf{Z} - \mathsf{Z}^*\mathsf{H}) = \xi[(\xi, \xi_1', \xi') + (\xi_1, \xi_2', \xi')] + \xi'[(\xi', \xi_2, \xi) + (\xi_2', \xi_1, \xi)].$$

Wird nun eine jede der vier durch die Klammern angedeuteten Determinanten nach den Elementen der ersten Spalte entwickelt, wobei die Unterdeterminanten mittels der schon wiederholt benutzten Relationen

$$\eta_1'\zeta' - \zeta_1'\eta' = n^{-1/3}\xi_1 - n^{1/3}\xi \quad \text{usw.}$$

umzuformen sind, so ergibt sich in der Tat:

$$3(\mathsf{H}^*\mathsf{Z} - \mathsf{Z}^*\mathsf{H}) = (\sum \xi_1\xi_2 - u)\xi + (\sum \xi_1'\xi_2' - v)\xi'.$$

---

[30]) Die Darstellung $x^* = 3\langle \mathsf{H}^*, \mathsf{Z}^*\rangle, \ldots$ gestattet diesen Schluß noch nicht; es könnten drei kongruente, aber durch Translationen unterschiedene Flächen vorliegen.

[31]) Es sei z. B. auf die Fußnote [27]) verwiesen.

Wir sprechen das hiermit gewonnene Schlußergebnis aus:

Die Transformation Θ_n *läßt sich in zwei asymptotische Transformationen zerlegen. Die Fläche* (x^*, y^*, z^*), *die den gemeinsamen zweiten Brennflächenmantel zweier je eine der beiden Biegungsflächen* (x, y, z) *und* (x_1, y_1, z_1) *berührender W-Kongruenzen bildet, erhält man, indem man die Biegungsfläche* (x, y, z) *noch einer zweiten (auf* ∞^1 *Weisen bestimmbaren*[32]) *Transformation* $\Theta_{\frac{1}{n}}$ *unterwirft, die zu der gegebenen* Θ_n *in harmonischer Beziehung steht. Ist* (x_2, y_2, z_2) *die so gewonnene dritte Biegungsfläche, so wird die Fläche* (x^*, y^*, z^*) *von dem Schnittpunkt entsprechender Tangentialebenen von* (x, y, z), (x_1, y_1, z_1) *und* (x_2, y_2, z_2) *beschrieben; gleichzeitig ist ihre Tangentialebene durch die entsprechenden Punkte der drei Biegungsflächen bestimmt.*

Berlin-Steglitz, im September 1923.

---

[32]) Auch aus dem allgemeinen Bianchischen Kompositionstheorem kann man folgern, daß die Zerlegung in zwei asymptotische Transformationen, wenn sie überhaupt gelingt, auf ∞^1 Weisen möglich ist.

(Eingegangen am 30. 9. 1923.)

Zur Theorie der topologischen Räume.

Von

Paul Alexandroff und Paul Urysohn † in Moskau.

Ein topologischer Raum $\mathfrak{R}$ entsteht, wenn wir in einer Menge E (die ganz abstrakt gegeben sein kann) gewisse Teilmengen, die *Umgebungen* ihrer sämtlichen Punkte, derart definieren, daß die bekannten vier Hausdorffschen Umgebungsaxiome[1]) damit zur Geltung gebracht werden. Ein topologischer Raum kann gleichzeitig durch verschiedene Umgebungssysteme definiert werden, die dann aber notwendig *gleichwertig*[2]) sind. Auch umgekehrt definieren, unserer Fassung nach, gleichwertige Umgebungssysteme stets einen einzigen topologischen Raum. Dieser Punkt scheint uns methodologisch von großer Wichtigkeit zu sein, wir dürfen aber hier nicht weiter auf ihn eingehen — wir wollen hier nur eine kurze Übersicht der Hauptergebnisse unserer Untersuchungen im Gebiete der allgemeinen Topologie angeben; für eine genaue Durchführung der Beweise sowie auch für die Konstruktion der zahlreichen Beispiele verweisen wir deshalb auf eine Arbeit, die bald in der Zeitschrift „Fundamenta Mathematicae" erscheinen soll.

1. Unter allen topologischen Räumen spielen, wie bekannt, die kompakten[3]) Räume eine besonders wichtige Rolle. Falls wir einen jeden der Relation[4])

$$|U(\xi)\cdot\mathfrak{M}| = |\mathfrak{M}|$$

genügenden Punkt ξ des Raumes $\mathfrak{R}$ als *vollständigen Häufungspunkt* der im Raume $\mathfrak{R}$ gelegenen Menge $\mathfrak{M}$ bezeichnen, besteht ersichtlich folgender

---

[1]) Hausdorff, Grundzüge der Mengenlehre, Kap. VII, S. 213.

[2]) Hausdorff, loc. cit., S. 260.

[3]) Hausdorff, loc. cit., S. 230; ein kompakter Raum ist natürlich in sich kompakt.

[4]) Wir bezeichnen durch $U(\xi)$ eine jede willkürlich gegebene Umgebung des Punktes ξ im Raume $\mathfrak{R}$, durch $A\cdot B\cdot C\ldots$ bzw. ΠA_n den Durchschnitt der Mengen $A, B, C, \ldots$ bzw. der Mengen A_n, durch $A+B+C\ldots$ bzw. ΣA_n die Vereinigungsmenge der (nicht notwendig elementefremden) Mengen $A, B, C, \ldots$ bzw. A_n; die Mächtigkeit der Menge $\mathfrak{M}$ soll stets mit $|\mathfrak{M}|$ bezeichnet werden.

Satz I_0. — *Damit der topologische Raum $\mathfrak{R}$ kompakt sei, ist eine jede (und folglich alle drei) der folgenden Bedingungen notwendig und hinreichend:*

A_0. *Eine jede abzählbare im Raume $\mathfrak{R}$ gelegene Menge $\mathfrak{M}$ besitzt daselbst einen vollständigen Häufungspunkt.*

B_0. *Eine abzählbare absteigende Folge (in $\mathfrak{R}$) abgeschlossener von Null verschiedener Mengen hat einen von Null verschiedenen Durchschnitt.*

C_0. *Ist der Raum $\mathfrak{R}$ in der Summe einer abzählbaren Menge von Gebieten enthalten, so ist er bereits in einer Summe von endlich vielen Gebieten dieser selben Menge enthalten.*

Nun läßt sich folgender Satz beweisen:

Satz I. *Folgende drei Eigenschaften A, B, C sind untereinander äquivalent (d. h. ein jeder topologischer Raum $\mathfrak{R}$, der irgendeine von diesen Eigenschaften hat, besitzt auch die beiden anderen):*

A. Eine jede im Raume $\mathfrak{R}$ gelegene unendliche Menge $\mathfrak{M}$ besitzt einen vollständigen Häufungspunkt.

B. Eine jede wohlgeordnete absteigende Menge[5]) *abgeschlossener von Null verschiedener Mengen hat einen von Null verschiedenen Durchschnitt.*

C. Ist der Raum $\mathfrak{R}$ in der Summe eines Systems (beliebiger Mächtigkeit) von Gebieten enthalten, so ist er bereits in einer Summe von endlich vielen Gebieten dieses Systems enthalten[5a]).

Die Äquivalenz ($B \sim C$) läßt sich durch formale Betrachtung von Produkt-, Summen- und Differenzbildungen ohne wesentliche Schwierigkeiten beweisen.

Indem wir mit dem Zeichen $\rightarrow$ das Wort „folgt" ersetzen, beweisen wir zunächst, daß $C \rightarrow A$; vorausgesetzt, es sei nicht der Fall, es existiere also eine unendliche Menge $\mathfrak{M}$ und eine gewisse Umgebung $U_0(x)$ eines jeden Punktes x des Raumes $\mathfrak{R}$, so daß

$$|\mathfrak{M}\cdot U_0(x)| < |\mathfrak{M}|$$

ist, so gelangen wir sofort zu einem Widerspruch: der ganze Raum ist nämlich zufolge der Eigenschaft C bereits in der Summe einer endlichen Anzahl von Gebieten $U_0(x)$ enthalten, und die Menge $\mathfrak{M}$ erscheint als Vereinigung endlich vieler Mengen $U_0(x)\cdot\mathfrak{M}$ von kleineren Mächtigkeiten, was offenbar unmöglich ist.

Indem wir durch A' bzw. B' die Eigenschaften bezeichnen, die durch Einschränkung der Aussagen A bzw. B auf Mengen $\mathfrak{M}$ bzw. $\mathfrak{S}$ von regulären[6]) Mächtigkeiten entstehen, ergibt sich sofort:

$$B' \rightarrow B.$$

---

[5]) Diese Menge wird nachher mit $\mathfrak{S}$ bezeichnet.

[5a]) In verschiedenen anderen Voraussetzungen (nicht in topologischen Räumen) sind analoge Sätze von Moore (Proc. Nat. Ac. Sciences **5**, 1919), Fréchet (Ann. Ec. Norm, 1921), Sierpiński (Fund. Math. 2) u. A. bewiesen worden.

[6]) Hausdorff, S. 130. $\aleph_\xi$ heißt regulär, wenn ω_ξ eine reguläre Ordnungszahl ist.

Der Satz wird demnach bewiesen sein, falls die beiden Formeln

$$B' \to A', \quad A' \to B'$$

richtig sind.

Es sei vorausgesetzt, A' sei nicht richtig, und $\mathfrak{M}$ eine Menge regulärer Mächtigkeit $\aleph_\tau$, die sich nach dem Typus[7]) ω_τ ordnen läßt; wenn x_α, $\alpha = 1, 2, 3, \ldots, < \omega_\tau$ sämtliche Punkte von $\mathfrak{M}$ sind, und $\mathfrak{M}_\lambda$ die Menge aller Punkte x_α, $\alpha \leqq \lambda < \omega_\tau$, so ergibt es sich, daß die wohlgeordnete (und zwar nach dem regulären Ordnungstypus ω_τ) Menge $\mathfrak{S}$ der abgeschlossenen Mengen[8])

$$F_\lambda = \overline{\mathfrak{M} - \mathfrak{M}_\lambda}$$

einen verschwindenden Durchschnitt hat, was der Eigenschaft B' widerspricht.

Wenn jetzt umgekehrt B' im Raume $\mathfrak{R}$ nicht erfüllt ist, dann gibt es eine wohlgeordnete abnehmende Menge $\mathfrak{S}$ abgeschlossener Mengen[9]):

$$F_1 \supset F_2 \supset F_3 \supset \ldots \supset F_n \supset \ldots \supset F_\alpha \supset \ldots \qquad (\alpha < \omega_\tau).$$

Man kann offenbar voraussetzen, daß außerdem $F_\alpha \neq F_{\alpha+1}$ ist. Dann wählt man für jedes α einen Punkt $x_\alpha \subset F_\alpha - F_{\alpha+1}$, und man beweist ohne große Schwierigkeit, daß die Menge aller Punkte x_α keinen vollständigen Häufungspunkt hat.

Aus dem soeben angedeuteten Beweise folgt außerdem, daß alle drei Eigenschaften A, B, C aus jeder einzelnen Eigenschaft A', B' und der analog gebildeten Eigenschaft C' folgen. Es liegt nahe, die Räume, in denen diese Eigenschaften stattfinden, besonders zu berücksichtigen und sie durch einen speziellen Namen auszuzeichnen; wir nennen sie also **bikompakt.**

2. Man kann den Satz I_0 wie folgt verallgemeinern: Man kann nämlich die Äquivalenz der Eigenschaften $A_\mathfrak{m}$, $B_\mathfrak{m}$, $C_\mathfrak{m}$, die aus den Eigenschaften A_0, B_0, C_0 entstehen, indem man das Wort „abzählbar" durch „von Mächtigkeit $\leqq \mathfrak{m}$" ersetzt, beweisen. Es ist natürlich, die dazu gehörigen Räume als initial kompakte Räume zu bezeichnen, und zwar: bis zu der Mächtigkeit $\mathfrak{m}$. *Die gewöhnlichen kompakten Räume erscheinen also als ein ganz willkürlicher Spezialfall.* Man kann auch eine gewisse finale Kompaktheit definieren, von der Mächtigkeit $\mathfrak{n}$ an. Das geschieht durch den folgenden Äquivalenzsatz:

---

[7]) Ebenda, S. 125.

[8]) Wir bezeichnen stets durch $\overline{\mathfrak{M}}$ die abgeschlossene Hülle der Menge $\mathfrak{M}$ (Hausdorff, S. 219—220 bezeichnet dieselbe Menge durch $\mathfrak{M}_\alpha$).

[9]) $\mathfrak{M} \supset M$ bedeutet, daß jeder Punkt der Menge M auch der Menge $\mathfrak{M}$ angehört (es kann insbesondere auch vorkommen, daß beide Mengen identisch sind); es kann auch sein, daß M (oder auch beide Mengen) nur einen Punkt enthält.

Es gilt

$$\begin{array}{ccc} A'_{(\mathfrak{n})} & \rightleftarrows & B'_{(\mathfrak{n})} \\ & \nwarrow\!\searrow \quad \swarrow\!\nearrow & \\ & C'_{(\mathfrak{n})} & \end{array},$$

wo $A'_{(\mathfrak{n})}$ und $B'_{(\mathfrak{n})}$ aus A' und B' entstehen, indem man die dazugehörigen Aussagen auf die Mengen $\mathfrak{M}$ und $\mathfrak{S}$ von regulären Mächtigkeiten $\geqq \mathfrak{n}$ einschränkt. $C'_{(\mathfrak{n})}$ gestaltet sich folgendermaßen:

$C'_{(\mathfrak{n})}$. Ist der Raum $\mathfrak{R}$ in der Summe einer Menge von regulärer Mächtigkeit $\geqq \mathfrak{n}$ von Gebieten enthalten, so ist er bereits in der Summe einer Menge von kleinerer Mächtigkeit dieser Gebiete enthalten.

Man sieht also, daß *die bikompakten Räume gleichzeitig initial- und final-kompakt sind, und zwar für jede* die Mächtigkeit der Menge aller Raumpunkte nicht überschreitende *unendliche Kardinalzahl.*

Beispiele. 1. Ein jeder kompakte metrische Raum ist auch bikompakt[10]). 2. Eine jede geordnete Menge läßt sich als topologischer Raum betrachten, indem man als $U(\xi)$ für einen jeden Punkt ξ die Mengen $J_{x,y}$ betrachtet, wo $J_{x,y}$ der Inbegriff aller zwischen x und y gelegenen Punkte ist und $x \prec \xi \prec y$. In diesem Sinne ist eine geordnete Menge, als Raum betrachtet, dann und nur dann bikompakt, wenn sie lückenlos ist und ein erstes sowie auch ein letztes Element besitzt. Insbesondere ist die wohlgeordnete Menge aller Ordnungszahlen $< \omega_\tau$ initial kompakt bis $\aleph_\tau$, die Menge aller Zahlen $\leqq \omega_\tau$ aber bikompakt.

3. Die bikompakten Räume ermöglichen den Aufbau einer ziemlich vollkommenen Theorie, deren Gedankengang in mehreren kleinen Abhandlungen kurz angezeigt sein soll. Hier kommen insbesondere die Fragen über die Stellung der bikompakten Räume unter den sämtlichen topologischen Räumen in Betracht. Wir definieren zuerst:

Ein topologischer Raum heißt ***absolut abgeschlossen****, falls er in jedem ihn umfassenden Raume R abgeschlossen ist*[11]).

Man kann auch sagen: Ein topologischer Raum $\mathfrak{R}$ ist dann und nur dann absolut abgeschlossen, wenn es unmöglich ist, ihm einen neuen Punkt ξ derart hinzuzufügen, daß die so entstandene Menge $\mathfrak{R} + \xi$

1. einen topologischen Raum, in dem ξ kein isolierter Punkt ist, bildet,

2. den ursprünglichen Raum $\mathfrak{R}$ mit sämtlichen Umgebungen seiner Punkte als Relativgebiet enthält.

---

[10]) Hausdorff, S. 272—274, der Satz folgt aus den Hausdorffschen Sätzen VI und X.

[11]) Der größere Raum R muß $\mathfrak{R}$ als *Raum* enthalten, d. h. die Relativumgebungen der Punkte von $\mathfrak{R}$ in R sollen den ursprünglichen Umgebungen derselben Punkte in $\mathfrak{R}$ gleichwertig sein. Man vergleiche wegen der Relativbegriffe Hausdorff, Kap. VII, § 6.

Es besteht folgender

Satz II. *Damit ein Raum $\mathfrak{R}$ absolut abgeschlossen sei, ist notwendig und hinreichend, daß aus jedem den Raum $\mathfrak{R}$ bedeckenden System von Gebieten eine endliche Anzahl derart ausgewählt werden kann, daß die betreffenden Gebiete nebst ihren Grenzpunkten den Raum vollständig ausfüllen*[12]).

Die Bedingung ist notwendig. Vorausgesetzt, sie sei nicht erfüllt; es existiere also ein System von Gebieten $\{G\}$, so daß für jedes endliche Aggregat

$$G_1, G_2, \ldots, G_n$$

dieser Gebiete die Relation

$$\mathfrak{G} = \mathfrak{R} - (\overline{G}_1 + \overline{G}_2 + \ldots + \overline{G}_n) \neq 0$$

besteht. Die $\mathfrak{G}$ sind wieder von Null verschiedene Gebiete. Wenn wir jetzt einen neuen Punkt ξ mittels der Umgebungen $U(\xi) = \xi + \mathfrak{G}$ einführen, erhalten wir einen neuen Raum $R = \mathfrak{R} + \xi$, in welchem $\mathfrak{R}$ eine nicht abgeschlossene Teilmenge ist; $\mathfrak{R}$ ist folglich kein absolut abgeschlossener Raum. Die Bedingung ist auch hinreichend. Vorausgesetzt, sie sei erfüllt und die Adjunktion eines Punktes ξ doch möglich. Indem wir $R = \mathfrak{R} + \xi$ betrachten, dürfen wir für einen jeden Punkt $x \subset \mathfrak{R}$ eine Umgebung $U_0(x)$ derart wählen, daß im Raume R $\overline{U}_0(x) \cdot \xi = 0$ ist. Nach unserer Voraussetzung ist eine endliche Menge der Gebiete $U_0^{(i)}$ derart angebbar, daß

$$\overline{U}_0^{(1)} + \overline{U}_0^{(2)} + \ldots + \overline{U}_0^{(n)} = \mathfrak{R}$$

ist. Das Gebiet $R - \sum_{i=1}^{n} \overline{U}_0^{(i)}$ enthält demnach nur einen einzigen Punkt ξ, der also isoliert in R ist, w. z. b. w.

Nun läßt sich folgender Satz leicht beweisen:

Satz III. *Es sei $\mathfrak{M}$ eine beliebige im absolut abgeschlossenen Raum $\mathfrak{R}$ gelegene unendliche Menge; es existiert dann stets ein Punkt ξ derart, daß*

$$|\overline{U}(\xi) \cdot \mathfrak{M}| = |\mathfrak{M}|$$

ist, wo $U(\xi)$ eine beliebige Umgebung des Punktes ξ ist.

Mit diesen Sätzen ist die Analogie zwischen den absolut abgeschlossenen und den bikompakten Räumen genügend deutlich geworden. Man könnte sie auch weiterführen, indem man z. B. ein Analogon der (B)-Eigenschaft gäbe; der Weg dazu scheint uns schon angedeutet zu sein. Man könnte sogar sagen, daß die absolut abgeschlossenen Räume *beinahe* bikompakt sind[13]);

---

[12]) Man vergleiche diesen Satz mit der (C)-Eigenschaft der bikompakten Räume.

[13]) Den Ausdruck verdanken wir Herrn Paul Bernays.

es gibt jedoch absolut abgeschlossene Räume, die nicht einmal kompakt sind (die auch dem II. Abzählbarkeitsaxiom genügen dürfen).

4. Um aus den absolut abgeschlossenen Räumen die bikompakten herauszubekommen, führen wir folgende Begriffe ein:

1. *Ein Punkt* ξ *heißt regulär im Raume* $\mathfrak{R}$, *falls für jedes* $U(\xi)$ *ein* $V(\xi)$ *derart existiert, daß* $U(\xi) \supset \overline{V}(\xi)$.

2. ***Ein nur reguläre Punkte enthaltender topologischer Raum heißt schlechthin regulär.***

Es läßt sich sofort beweisen, daß, damit ein Raum regulär sei, es notwendig und hinreichend ist, daß ein beliebiger Punkt ξ von einer ihn nicht enthaltenden abgeschlossenen Menge F mittels zweier punktfremder Gebiete U und G abgetrennt werden kann, derart, daß $U \supset \xi$, $G \supset F$, $U \cdot G = 0$.

Es liegt nun nahe, diese letzte Eigenschaft noch folgendermaßen zu verschärfen:

3. *Ein Raum heißt normal, falls für je zwei punktfremde abgeschlossene Mengen* F_1 *und* F_2 *zwei ebenfalls punktfremde Gebiete* G_1 *und* G_2 *existieren, derart, daß* $G_1 \supset F_1$ *und* $G_2 \supset F_2$ (es läßt sich beweisen, daß in normalen Räumen G_1 und G_2 so gewählt werden können, daß auch $\overline{G}_1 \cdot \overline{G}_2$ Null ist[14])).

Ein normaler topologischer Raum ist *a fortiori* regulär.

Es ergibt sich nun

Satz IV. *Ein jeder bikompakte topologische Raum ist normal*[15]).

Der Satz läßt sich mittels zweimaliger Anwendung des sog. Borel-Lebesgueschen Satzes (Eigenschaft C der bikompakten Räume) ohne Mühe beweisen.

Auf den Sätzen III und IV fußend kann man endlich den folgenden Fundamentalsatz beweisen:

Satz V. *Damit ein topologischer Raum bikompakt sei, ist es notwendig und hinreichend, daß er regulär und absolut abgeschlossen sei.*

Durch diese Sätze scheint uns die Definition der regulären Räume als einer besonders anschaulichen und naturgemäßen Klasse von Räumen gerechtfertigt zu sein; wir erwähnen nur noch, daß in den irregulären

---

[14]) Wie uns vor einigen Tagen bekannt geworden ist, benutzt Herr Tietze in seiner Arbeit „Beiträge zur allgemeinen Topologie" (Math. Annalen 88) analoge Begriffsbildungen. — Der während der Drucklegung dieser Arbeit erschienene II. Teil der Tietzeschen „Beiträge" hat auch manche Berührungspunkte mit unseren Untersuchungen.

[15]) Der Satz besteht allgemein nicht in kompakten Räumen; vielmehr sind auch irreguläre kompakte Räume von uns konstruiert worden; ein jeder kompakte, dem I. Abzählbarkeitsaxiome genügende topologische Raum ist aber regulär.

Räumen sehr verschiedene, zum Teil auch sehr eigentümliche Singularitäten vorhanden sein können; insbesondere gibt es z. B. zusammenhängende abzählbare Räume usw. Überhaupt scheinen uns die topologischen Eigenschaften der irregulären Räume in einer so unregelmäßigen Gestalt vorzukommen, daß sie einer noch einigermaßen einfachen Theorie kaum fähig sind[16]).

5. Der Begriff der absoluten Abgeschlossenheit eines Raumes kann auch weitergeführt werden. In dieser Hinsicht denken wir uns alle die Punkte des Raumes $\mathfrak{R}$, die durch eine gewisse Eigenschaft $\mathfrak{E}$ charakterisiert sind; wir nennen diese Punkte etwa die $\mathfrak{E}$-Punkte des Raumes $\mathfrak{R}$. Man kann z. B. als Eigenschaft $\mathfrak{E}$ die Eigenschaft eines Punktes, regulär zu sein, wählen. Wir nennen jetzt einen Raum $\mathfrak{R}$ abgeschlossen in bezug auf seine sämtlichen $\mathfrak{E}$-Punkte (oder auch einfach „$\mathfrak{E}$-*abgeschlossen*"), falls es unmöglich ist, dem Raume $\mathfrak{R}$ einen Punkt ξ derart hinzuzufügen, daß er im erweiterten Raume $\mathfrak{R} + \xi = R$ die Eigenschaft $\mathfrak{E}$ besitze und nicht isoliert sei.

Nach dieser Definition ist ein absolut abgeschlossener Raum in bezug auf alle seine nicht isolierten Punkte abgeschlossen. Man bemerke noch, daß die Menge der sämtlichen $\mathfrak{E}$-Punkte eines $\mathfrak{E}$-abgeschlossenen Raumes eine leere Menge sein kann, wie man sich durch ganz elementare Beispiele sofort überzeugt.

Wenn wir jetzt definieren

1. Eigenschaft $(\varkappa)$ eines Punktes ξ im Raume $\mathfrak{R}$ = es existiert eine abzählbare gegen den Punkt ξ konvergierende im Raume $\mathfrak{R}$ gelegene Punktmenge;

2. Eigenschaft (ι) = die Menge aller Umgebungen des Punktes ξ im Raume $\mathfrak{R}$ ist einer abzählbaren Umgebungsmenge gleichwertig;

3. Eigenschaft (δ) = der Punkt ξ ist der Durchschnitt einer abzählbaren Menge von Gebieten,

so überzeugen wir uns zuerst durch Beispiele, daß die einzige Abhängigkeit, die zwischen den drei Eigenschaften $(\varkappa)$, (ι), (δ) besteht, darin liegt, daß die beiden Eigenschaften $(\varkappa)$ und (δ) aus (ι) folgen; $(\varkappa)$ und (δ) sind untereinander unabhängig; (ι) braucht sogar nicht aus den beiden Eigenschaften $(\varkappa)$ und (δ) zu folgen[17]).

Es besteht nun der leicht beweisbare Satz:

---

[16]) Wir verweisen nochmals auf unsere in den „Fundamenta Mathematicae" bald erscheinende ausführliche Darstellung der ganzen Theorie.

[17]) Über die in bikompakten Räumen vorhandenen Verhältnisse siehe P. Alexandroff, „Über die Struktur der bikompakten topologischen Räume" (dieser Band, S. 267–274).

Satz VI. *Damit ein regulärer topologischer Raum kompakt sei, ist es notwendig und hinreichend, daß er* ($\varkappa$)*-abgeschlossen sei*[18]).

Dagegen:

1. Es gibt absolut abgeschlossene nichtkompakte (natürlich irreguläre) Räume, die sogar dem II. Abzählbarkeitsaxiome genügen.

2. Es gibt kompakte, in bezug auf *sämtliche* (δ)-Punkte nicht abgeschlossene reguläre Räume; wohl aber

2'. Ein jeder kompakte Raum ist abgeschlossen in bezug auf seine *regulären* (δ)-Punkte.

2''. Es gibt nichtkompakte, in bezug auf sämtliche regulären (δ)-Punkte abgeschlossene reguläre Räume.

3. Es gibt nichtkompakte, (ι)-abgeschlossene reguläre Räume.

Es ergibt sich ferner:

Satz VII. *Damit ein normaler Raum kompakt sei, ist es notwendig und hinreichend, daß er* (ι)*-abgeschlossen sei.*

Der Satz VII gilt sogar in einer allgemeineren Form; es genügt nämlich, nur vorauszusetzen, daß eine jede abgeschlossene Menge F von jeder abzählbaren, divergenten Menge D (die zu F fremd ist) mittels zueinander fremder, die Mengen F und D enthaltender Gebiete G_F, G_D getrennt werden kann; falls diese Bedingung im Raume $\mathfrak{R}$ erfüllt ist, so ist der ganze Raum nicht (ι)-abgeschlossen, soweit er nicht kompakt ist.

Der springende Punkt beim Beweise des obigen Satzes ist der, daß man für irgendeine divergente abzählbare Punktmenge

$$x_1, x_2, \ldots, x_n, \ldots$$

gewisse Umgebungen

$$\text{(a)} \qquad U_0(x_1),\ U_0(x_2), \ldots, U_0(x_n), \ldots$$

konstruiert, die den Gleichungen genügen:

$$\overline{U}_0(x_i) \cdot \overline{U}_0(x_k) = 0$$

$$\overline{\sum_{i=1}^{\infty} U_0(x_i)} = \sum_{i=1}^{\infty} \overline{U}_0(x_i),$$

und dies geschieht auf Grund der erwähnten Trennungsvoraussetzungen ohne große Mühe.

[18]) Nur die eine Hälfte dieses Satzes gilt auch für irreguläre Räume, und zwar ist ein jeder kompakte topologische Raum offenbar ($\varkappa$)-vollständig (a fortiori auch (ι)-vollständig).

Nachdem die Konstruktion der Umgebungsmenge (a) gelungen ist, fügt man dem gegebenen Raume $\mathfrak{R}$ einen neuen Punkt ξ mittels der Umgebungserklärung

$$U_n(\xi) = \xi + \sum_{i=n+1}^{\infty} U_0(x_i)$$

hinzu.

Man kann außerdem diese Konstruktion auch so einrichten, daß der erweiterte Raum $\mathfrak{R} + \xi = R$ regulär ist (sogar denselben Trennungsbedingungen wie der ursprüngliche Raum $\mathfrak{R}$ genügt).

Die wichtige und interessante Frage, ob das auch immer der Fall ist, d. h. ob ein jeder reguläre nicht absolut abgeschlossene topologische Raum sich zu einem ebenfalls regulären Raume durch Hinzufügung eines nicht isolierten Punktes erweitern läßt, bleibt unentschieden. Wir werden an einer anderen Stelle noch andere Spezialfälle dieses Problems zur Lösung bringen.

Göttingen, den 26. Juni 1923.

P. S. Die Hauptergebnisse dieser Arbeit sind im März und Juni 1922 der Moskauer Mathematischen Gesellschaft und neuerdings der Göttinger Mathematischen Gesellschaft vorgetragen worden.

Während der Drucklegung dieses Heftes hat die Redaktion die erschütternde Nachricht erhalten, daß Paul Urysohn am 17. August im Alter von 26 Jahren durch einen Unglücksfall um das Leben gekommen ist.

(Eingegangen am 1. 8. 1923.)

Über die Struktur der bikompakten topologischen Räume.

Von

Paul Alexandroff in Moskau.

---

1. In einer gemeinsam mit Herrn Urysohn verfaßten Arbeit[1]) habe ich den Begriff der bikompakten topologischen Räume bereits eingeführt; daselbst ist auch die Stellung dieser Räume zu anderen topologischen Räumen gewissermaßen erläutert worden. In dem vorliegenden Aufsatze möchte ich in den Aufbau der bikompakten Räume tiefer eindringen und gewisse Struktureigenschaften derselben auseinandersetzen. Insbesondere möchte ich auf die Struktur der bikompakten Räume *im Kleinen*, d. h. in einer gewissen Umgebung eines jeden Raumpunktes, aufmerksam machen. Dazu fange ich mit folgendem elementaren Beispiele an.

Man denke an den aus allen Ordnungszahlen $\alpha \leqq \omega_1$ [2]) gemäß der Vorschrift $\mathfrak{H}$, § 2, Beisp. 2 gebildeten topologischen Raum $\mathfrak{R}$ (d. h. wo

$$\underset{0 \leqq \alpha < \Omega}{U} (\alpha+1) = (\alpha+1); \qquad \underset{\substack{\alpha = \lim \alpha_n \\ n \to \infty}}{U_\lambda} (\alpha) = \sum_{\lambda < \beta \leqq \alpha} (\beta)$$

die Umgebungen der den transfiniten Zahlen 1-ter bzw. 2-ter Art entsprechenden Punkte $(\alpha+1)$ bzw. (α) sind). Wir haben nun folgende Eigenschaften der zwei Punkte ω_0 und ω_1 im Raume $\mathfrak{R}$

(ω_0)	(ω_1)
1. Es gibt im Raume $\mathfrak{R}$ eine Gebietsmenge von der Mächtigkeit $\aleph_0$, deren Durchschnitt aus dem einzigen Punkte (ω_0) besteht.	1. Es gibt im Raume $\mathfrak{R}$ eine Gebietsmenge von der Mächtigkeit $\aleph_1$, deren Durchschnitt aus dem einzigen Punkte (ω_1) besteht.
2. Es gibt eine Umgebungsmenge von der Mächtigkeit $\aleph_0$, die den Punkt (ω_0) im Raume $\mathfrak{R}$ definiert.	2. Es gibt eine Umgebungsmenge von der Mächtigkeit $\aleph_1$, die den Punkt (ω_1) im Raume $\mathfrak{R}$ definiert.

---

[1]) „Zur Theorie der topologischen Räume", dieser Band, S. 258. Diese Arbeit wird in der Folge durch ein $\mathfrak{H}$ bezeichnet; die dort enthaltenen Ergebnisse und Bezeichnungen werden als bekannt vorausgesetzt.

[2]) ω_1 ist, nach Hausdorffs Bezeichnung, die erste Ordnungszahl von der Mächtigkeit $\aleph_1$.

(ω_0)

3. Es gibt eine Punktmenge E_0 von der Mächtigkeit $\aleph_0$, so daß
$$|U(\omega_0)\cdot E_0| > |(\mathfrak{R}-U(\omega_0))\cdot E_0|$$
ist ($U(\omega_0)$ ist eine beliebige Umgebung von (ω_0)).

(ω_1)

3. Es gibt eine Punktmenge E_1 von der Mächtigkeit $\aleph_1$, so daß
$$|U(\omega_1)\cdot E_1| > |(\mathfrak{R}-U(\omega_1))\cdot E_1|$$
ist ($U(\omega_1)$ ist eine beliebige Umgebung von (ω_1)).[3]

2. Nun liegen folgende Definitionen auf der Hand:

Definition 1. *Wir sagen, daß eine Menge $\mathfrak{M}$ im Raume $\mathfrak{R}$ gegen den Punkt ξ strömt, wenn für jede beliebige Umgebung $U(\xi)$ des Punktes ξ die Relation*
$$|U(\xi)\cdot\mathfrak{M}| > |(\mathfrak{R}-U(\xi))\cdot\mathfrak{M}|$$
zutrifft. In Zeichen „lim“ $\mathfrak{M} = \xi$.

Man sieht sofort:

1. Eine Menge kann nicht mehr als einen Strömungspunkt besitzen.

2. Eine abzählbare Menge strömt einem Punkte dann und nur dann zu, wenn sie zu diesem Punkte im Hausdorffschen Sinne[4]) konvergiert.

Definition 2. *Man bezeichnet als Konvergenzcharakter des Raumes $\mathfrak{R}$ im Punkte x die kleinste unendliche Kardinalzahl $\varkappa_{\mathfrak{R}}(x)$, welche die Eigenschaft besitzt, daß es in $\mathfrak{R}$ eine gegen den Punkt x strömende Menge $\mathfrak{M}$ von der Mächtigkeit $|\mathfrak{M}| = \varkappa_{\mathfrak{R}}(x)$ gibt.*

Anmerkung. Von den isolierten Punkten sagen wir, sie besitzen den Konvergenzcharakter 1.

Nun ist es besonders hervorzuheben, daß ein topologischer Raum in einem gegebenen Punkte überhaupt keinen Konvergenzcharakter zu besitzen braucht, in dem Sinne, daß der Punkt weder isoliert ist, noch eine zu ihm strömende unendliche Menge besitzt. Es ist sogar sehr leicht, Beispiele von Räumen anzugeben, in denen keine unendliche Menge einen Strömungspunkt besitzt. Im Gegenteil dazu hat ein jeder Punkt eines dem I. Abzählbarkeitsaxiome[5]) genügenden Raumes einen wohldefinierten Konvergenzcharakter, und zwar vom Betrage $\aleph_0$ oder 1. Als weitere Beispiele von Räumen, die sich in jedem Punkte mit einem bestimmten Konvergenzcharakter erweisen, können die aus den geordneten Mengen entstehenden topologischen Räume dienen.

---

[3]) Vgl. für die Bezeichnungen $\mathfrak{H}$ Fußnoten [4]), [5]), [9]). Insbesondere bedeutet $|\mathfrak{M}|$ = Mächtigkeit der Menge $\mathfrak{M}$, $|\mathfrak{R}|$ die Mächtigkeit der Menge aller Punkte des gegebenen Raumes $\mathfrak{R}$.

[4]) Hausdorff, Grundzüge der Mengenlehre, S. 232.

[5]) Hausdorff, loc. cit. S. 263. Insbesondere genügt jeder metrische Raum diesem Axiome.

Ein vollständiges Bild der Lage eines Punktes im Raume erhält man aber nur mit Hilfe weiterer Definitionen:

Definition 3. *Man bezeichnet als* ***Durchschnittscharakter*** *des Raumes* $\mathfrak{R}$ *im Punkte* x *die kleinste Kardinalzahl* $\psi_{\mathfrak{R}}(x)$, *welche die Eigenschaft hat, daß es in* $\mathfrak{R}$ *eine Menge* $\mathfrak{S}$, $|\mathfrak{S}| = \psi_{\mathfrak{R}}(x)$, *von Gebieten gibt, deren Durchschnitt aus dem einzigen Punkte* x *besteht.*

Definition 4. *Wir nennen schlechthin* ***Charakter*** $\chi_{\mathfrak{R}}(x)$ *des Raumes* $\mathfrak{R}$ *im Punkte* x *die kleinste Mächtigkeit einer den Punkt* x *im Raume* $\mathfrak{R}$ *definierenden Umgebungsmenge* (d. h. einer Menge, in der zu jedem den Punkt x enthaltendem Gebiete G eine in G enthaltene Umgebung von x vorhanden ist).

Sowohl der Charakter als auch der Durchschnittscharakter sind für jeden Punkt eines beliebigen Raumes wohldefinierte Kardinalzahlen, die z. B. für sämtliche nicht isolierten Punkte eines dem I. Abzählbarkeitsaxiome genügenden Raumes den Wert $\aleph_0$ haben.

Es ergibt sich unmittelbar aus den Definitionen, daß stets $\psi_{\mathfrak{R}}(x) \leqq \chi_{\mathfrak{R}}(x)$ ist; es braucht aber keineswegs $\psi_{\mathfrak{R}}(x) = \chi_{\mathfrak{R}}(x)$ zu sein. Es können auch alle drei Kardinalzahlen $\varkappa_{\mathfrak{R}}(x)$, $\psi_{\mathfrak{R}}(x)$ und $\chi_{\mathfrak{R}}(x)$ verschieden sein; daneben können auch zwei beliebige untereinander gleich, von der dritten aber verschieden sein; endlich kann bei $\psi = \chi$ ebensogut wie bei $\psi < \chi$ die dritte Zahl $\varkappa$ sowohl definiert als auch nicht definiert sein[6]).

3. Nun wird die Lage eines Punktes in einem bikompakten Raume $\mathfrak{R}$ durch folgende zwei Sätze charakterisiert:

Satz I. *Ein bikompakter topologischer Raum genügt in jedem seiner Punkte der Gleichung* $\psi_{\mathfrak{R}}(x) = \chi_{\mathfrak{R}}(x)$.

Satz II. *Der Konvergenzcharakter eines bikompakten Raumes ist in jedem seiner Punkte wohldefiniert, und zwar ist stets* $\varkappa_{\mathfrak{R}}(x) \leqq \chi_{\mathfrak{R}}(x)$.

Der Beweis des Satzes I geschieht mit Hilfe der naheliegenden Hilfssätze:

Hilfssatz 1. Damit ein Raum bikompakt sei, ist folgende Bedingung notwendig und hinreichend:

$\mathfrak{S}$ sei ein beliebiges System von abgeschlossenen Mengen F, von denen je endlich viele einen nicht leeren Durchschnitt haben; dann ist auch der Durchschnitt aller Mengen F von Null verschieden.

---

[6]) Ich verweise wegen der Beispiele für alle diese Möglichkeiten auf meine in den „Fundamenta Mathematicae" bald erscheinende Arbeit „Sur les espaces localement compacts".

Es ist noch zu bemerken, daß für eine reguläre Mächtigkeit $\mathfrak{m} = |\mathfrak{M}|$ es unmöglich ist, daß, falls $x =$ „lim" $\mathfrak{M}$ ist, $\chi_{\mathfrak{R}}(x) < \mathfrak{m}$ sei; wohl aber gibt es Räume, wo $\chi_{\mathfrak{R}}(x) < \varkappa_{\mathfrak{R}}(x)$ und die letzte Kardinalzahl irregulär ist (eine Kardinalzahl heißt rirregulär, wenn die zugehörige Anfangszahl einer kleineren konfinal ist (Hausdorff, loc. cit. S. 130)).

Hilfssatz 2. Aus jeder den Punkt x im Raume $\mathfrak{R}$ definierenden Umgebungsmenge $\Sigma^{(x)}$ kann man eine Teilmenge $\Sigma_0^{(x)}$ von der Mächtigkeit $\psi_{\mathfrak{R}}(x)$ derart herausgreifen, daß der Durchschnitt aller betreffenden Umgebungen nur den Punkt x enthält.

Es sei nun $\Sigma^{(x)}$ die Menge aller Umgebungen des Punktes x im Raume $\mathfrak{R}$, $\Sigma_0^{(x)}$ die infolge des Hilfssatzes 2 existierende Teilmenge von der Mächtigkeit $\psi_{\mathfrak{R}}(x)$. Es seien ferner $V(x)$ die Umgebungen des Systems $\Sigma_0^{(x)}$. Wir ergänzen nun das System $\Sigma_0^{(x)}$ folgendermaßen:

1. Für ein jedes $V(x)$ wählen wir ein gewisses $U_V(x)$, das der Relation $\bar{U}_V(x) \subset V(x)$ [7]) genügt. Die Gesamtheit aller $V(x)$ und $U_V(x)$ bezeichnen wir durch $\Sigma_1^{(x)}$.

2. Wir erhalten ein Umgebungssystem $\Sigma_2^{(x)}$, indem wir alle Durchschnittsmengen von je endlich vielen Mengen des Systems $\Sigma_1^{(x)}$ betrachten.

Das System $\Sigma_2^{(x)}$ besitzt offenbar die Mächtigkeit $\psi_{\mathfrak{R}}(x)$. Nun läßt es sich mit Hilfe des Hilfssatzes 1 beweisen, daß die beiden Umgebungssysteme $\Sigma_2^{(x)}$ und $\Sigma^{(x)}$ gleichwertig sind, und damit ist der Satz I als richtig erwiesen.

Bevor wir zum Beweise des Satzes II schreiten, machen wir noch folgende Bemerkungen:

Korollar 1. *Wenn in einem bikompakten Raume eine jede abgeschlossene Menge ein Durchschnitt von höchstens abzählbar vielen Gebieten ist* (= *ein jedes F ein G_δ ist*[8])), *so genügt der Raum dem* I. *Abzählbarkeitsaxiome.*

Dieser Satz gilt allgemeiner in *regulären* kompakten Räumen, aber nicht in allgemeinen topologischen (auch kompakten!) Räumen. Die Umkehrung des Satzes ist dagegen überhaupt falsch.

Korollar 2. $\chi_{\mathfrak{R}}(\xi)$ *ist stets* $\leq |U(\xi)|$, wo $|U(\xi)|$ die Mächtigkeit der Menge aller Punkte einer beliebigen Umgebung des Punktes ξ im bikompakten Raume $\mathfrak{R}$ bezeichnet[9]).

In kompakten Räumen gilt nun der folgende Satz:

I_0. *Falls in einem regulären Punkte ein kompakter Raum den Durchschnittscharakter $\aleph_0$ besitzt, so ist daselbst sein Charakter auch $\aleph_0$.*

Wie man durch Gegenbeispiele zeigt, ist dieser Satz keiner Verschärfung fähig.

---

[7]) Dies ist zufolge der Regularitätseigenschaft der bikompakten Räume stets möglich (man vgl. dafür $\mathfrak{H}$, z. B. den Satz IV). (Ein Raum heißt regulär, falls ein jedes $U(x)$ ein gewisses $\bar{V}(x)$ enthält.)

[8]) Hausdorff, loc. cit. S. 23 u. 304.

[9]) Dagegen sind von Herrn Urysohn topologische Räume konstruiert worden, in denen die Charaktere einzelner Punkte die Mächtigkeit der Menge aller Raumpunkte übertreffen.

Wir schreiten nun zur kurzen Übersicht des Beweises von Satz II.

Es sei eine vollständige Umgebungsmenge Σ^x von der Mächtigkeit $\chi_{\mathfrak{R}}(x)$ gegeben. Auf dem Zermeloschen Wohlordnungssatze fußend, kann man sich Σ^x als eine wohlgeordnete Menge

$$V_1(x),\ V_2(x), \ldots, V_n(x), \ldots, V_\alpha(x) \ldots$$

vom Ordnungstypus Ω_χ denken, wo Ω_χ die kleinste Ordnungszahl von der Mächtigkeit $\chi_{\mathfrak{R}}(x) = \psi_{\mathfrak{R}}(x)$ ist.

Es sei $F_1 = \overline{V}_1$ gesetzt. Nehmen wir an, F_α sei bereits konstruiert; es sei dann $\tau(\alpha)$ die erste Ordnungszahl von der Eigenschaft, daß $F_\alpha - \overline{V}_{\tau(\alpha)} \cdot F_\alpha \neq 0$ ist. Man setzt dann $F_{\alpha+1} = F_\alpha \cdot \overline{V}_{\tau(\alpha)}$. Falls alle F_α, $\alpha < \lambda$, $\lambda \neq \mu + 1$, schon konstruiert sind, setzen wir $F_\lambda = \prod\limits_{\alpha<\lambda} F_\alpha$. Das Verfahren kann nur dann abbrechen, wenn F_λ nur den einzigen Punkt ξ enthält. Wir bekommen so eine wohlgeordnete abnehmende Menge von paarweise verschiedenen abgeschlossenen Mengen. Man überzeugt sich leicht, daß jedenfalls der Ordnungstypus Θ dieser wohlgeordneten Menge der Ordnungszahl Ω_χ konfinal ist, und das genügt, um ohne Schwierigkeit zu beweisen, daß die Menge aller Punkte x_α, $1 \leqq \alpha < \Theta$, die willkürlich aus den Mengen $F_\alpha - F_{\alpha+1}$ gewählt worden sind, eine zu dem Punkte x strömende unendliche Menge von der Mächtigkeit $|\Theta| \leqq \chi_{\mathfrak{R}}(x)$ ist.

Damit ist der Beweis des Satzes II erbracht.

4. Wir wollen jetzt den Zusammenhang zwischen den Eigenschaften im Kleinen und gewissen Eigenschaften, die dem Gesamtraume zukommen, insbesondere den Mächtigkeitseigenschaften der bikompakten Räume, näher betrachten.

Zuerst folgt der

Satz III. *Jede in einem bikompakten Raume liegende perfekte Menge hat die Mächtigkeit* $\geqq \mathfrak{c}$

unmittelbar aus der Regularität der bikompakten Räume [10]).

Es entsteht nun die Frage nach dem Vorhandensein von perfekten Teilmengen in bikompakten Räumen. Diese Frage läßt sich beantworten, und zwar mit Hilfe folgender Definition:

Defintion 5. *Ein Punkt ξ des topologischen Raumes $\mathfrak{R}$ heißt ein Staupunkt, wenn, wie groß die Kardinalzahl* $\mathfrak{m} < |\mathfrak{R}|$ *auch gewählt sei, in jeder Umgebung $U(\xi)$ Punkte x mit* $\chi_{\mathfrak{R}}(x) > \mathfrak{m}$ *vorkommen.*

Zufolge dieser Definition ist die Menge Ξ aller Staupunkte des Raumes $\mathfrak{R}$ eine abgeschlossene Menge; falls ξ in Ξ isoliert ist, heißt er ein isolierter Staupunkt. Jetzt formulieren wir den

---

[10]) $\mathfrak{H}$, Satz IV.

Satz IV. *Jeder bikompakte Raum ohne perfekte Teilmenge enthält stets einen isolierten Staupunkt.*

Wenn die Mächtigkeit $|\mathfrak{R}|$ der Menge aller Raumpunkte eine reguläre Mächtigkeit ist[11]), kann man den Satz noch verschärfen, indem man zeigt, daß in diesem Falle der Raum stets einen Punkt ξ, dessen Charakter $\chi_{\mathfrak{R}}(\xi) = |\mathfrak{R}|$ ist, enthält.

Der Fall der regulären Mächtigkeit $|\mathfrak{R}|$ wird sofort durch den folgenden leicht beweisbaren Hilfssatz erledigt:

Hilfssatz. Wenn $|\mathfrak{R}|$ regulär ist, und ξ ein isolierter Punkt der abgeschlossenen, aus sämtlichen vollständigen Häufungspunkten des gesamten Raumes $\mathfrak{R}$ bestehenden Menge Φ ist, so ist $\chi_{\mathfrak{R}}(\xi) = |\mathfrak{R}|$.

Der allgemeine Fall läßt sich folgendermaßen untersuchen.

Es sei $|\mathfrak{R}|$ irregulär und ξ ein vollständiger Häufungspunkt von $\mathfrak{R}$, der in der Menge Φ aller derartigen Punkte isoliert ist. Ich behaupte, ξ sei ein Staupunkt. Vorausgesetzt $\chi_{\mathfrak{R}}(\xi) < |\mathfrak{R}|$, so erhält man für jede Umgebung $U(\xi)$ zwei andere Umgebungen $U_0(\xi)$ und $U_1(\xi)$, so daß $U_0(\xi) \subset U_1(\xi) \subset \overline{U}_1(\xi) \subset U(\xi)$ und außerdem $|\overline{U}_1(\xi) - U_0(\xi)| > \mathfrak{r} > \mathfrak{m}$ ist, wo $\mathfrak{r}$ und $\mathfrak{m}$ beliebige Kardinalzahlen unter $|\mathfrak{R}|$ sind, und $\mathfrak{r}$ als regulär vorausgesetzt werden darf. Man greift nun eine Teilmenge E, $|E| = \mathfrak{r}$ aus der abgeschlossenen Menge $\overline{U}_1(\xi) - U_0(\xi)$ heraus; da der Raum keinen perfekten Bestandteil hat, so enthält die abgeschlossene Menge F aller vollständigen Häufungspunkte der Menge E sicher einen isolierten Punkt $x \subset F \subset \overline{U}_1(\xi) - U_0(\xi) \subset U(\xi)$, und es ist leicht einzusehen, daß $\chi_{\mathfrak{R}}(x) \geqq \mathfrak{r} > \mathfrak{m}$ ist. Damit ist aber der Satz IV offenbar bewiesen.

Aus dem soeben gewonnenen Ergebnisse folgt unmittelbar der

Satz V. *Jeder das 1. Abzählbarkeitsaxiom erfüllende bikompakte topologische Raum ist entweder abzählbar oder mit einer perfekten Teilmenge versehen; im letzteren Falle ist die Mächtigkeit des Raumes* $\geqq \mathfrak{c}$.

Es ist besonders bemerkenswert, daß unter der Voraussetzung des Satzes V die Zerspaltung des Raumes in einen perfekten Kern und eine abzählbare Punktmenge im allgemeinen nicht möglich zu sein braucht; vielmehr gibt es bikompakte, dem I. Abzählbarkeitsaxiome genügende Räume, die eine unabzählbare (und zwar von der Mächtigkeit $\mathfrak{c}$) Menge von isolierten Punkten enthalten[12]). Ich bemerke noch, daß die Sätze

[11]) Eine Mächtigkeit $\mathfrak{m} = \aleph_\tau$ heißt bekanntlich regulär, wenn die zugehörige Anfangszahl ω_τ keiner kleineren Ordnungszahl konfinal ist. Man vergl. Hausdorff, loc. cit. S. 130.

[12]) Die Frage, ob es bikompakte, dem I. Abzählbarkeitsaxiome genügende Räume von der Mächtigkeit $> \mathfrak{c}$ geben könne, bleibt bisher ungelöst. Herr Urysohn hat bewiesen, daß diese Frage sich verneinen läßt in bikompakten Räumen, in denen eine jede abgeschlossene Menge Durchschnitt von abzählbar-vielen Gebieten ist (= ein

IV und V in kompakten (sogar regulären, aber nicht bikompakten) Räumen allgemein falsch sind.

5. Ich will nun jetzt einen viel allgemeineren Mächtigkeitssatz aussprechen, der, soviel mir bekannt ist, überhaupt die allgemeinste Behauptung über Mächtigkeiten von Punktmengen enthält, die wir, nach dem heutigen Stande der Wissenschaft, aufstellen können. Es handelt sich zunächst um die sogenannten Borelschen Mengen, deren kürzeste Definition die folgende, von Sierpiński herrührende ist:

Wir nennen ein System von Mengen ein T-System, wenn Summe und Durchschnitt von je höchstens abzählbar-vielen Mengen des Systems wieder dem Systeme angehören. Dann ist der Inbegriff aller Borelschen Mengen definitionsgemäß das kleinste (d. h. in jedem anderen enthaltene) T-System, welches alle abgeschlossenen Mengen und alle Gebiete des gegebenen Raumes enthält.

Es war seit langer Zeit eine fast allgemeine Meinung, daß man mit den üblichen Mitteln der Analysis nicht über die Borelschen Mengen hinausgelangen könne. Erst im Jahre 1917 zeigte Souslin, daß *die Borelschen Mengen nur ein Spezialfall einer viel weiter gehenden Klasse der sogenannten* (A)-*Mengen sind*[13]), die sich folgendermaßen definieren

jedes F ein G_δ ist, Hausdorff, loc. cit. S. 23 und 304, 305). In denselben Räumen ist auch der Cantor-Bendixsonsche Satz gültig, d. h. eine jede abgeschlossene unendliche Menge ist entweder abzählbar oder Summe einer perfekten (von Null verschiedenen) und einer höchstens abzählbaren Menge. Vgl. Paul Alexandroff und Paul Urysohn, „Sur les espaces topologiques compacts" (wird in den „Fundamenta Mathematicae" erscheinen).

[13]) Der erste Beweis, daß eine *jede Borelsche Menge eine* (A)-*Menge ist*, ist zuerst vollständig durchgeführt, aber nicht formuliert worden in meiner Comptes-Rendus-Note „*Sur la puissance des ensembles mesurables B*" (C. R., 28 février 1916). Diese Tatsache ist gerade der springende Punkt des Beweises des Mächtigkeitssatzes, der dort gegeben ist; der Mächtigkeitsbeweis läßt sich wörtlich auf allgemeine (A)-Mengen übertragen. Souslin hat im Januar 1917 (ebenfalls in den Comptes Rendus) den Begriff der allgemeinen (A)-Mengen zuerst explizite formuliert und *ein Beispiel einer* (A)-*Menge, die keine* (B)-*Menge ist*, gegeben. Herr Lusin bewies danach, daß *die Wertmenge einer jeden analytisch-darstellbaren Funktion* (= einer Baireschen Funktion) *stets eine* (A)-*Menge, aber nicht allgemein eine* (B)-*Menge ist.* Ich erwähne aus der ganzen tiefgehenden Soulin-Lusinschen Theorie noch den Satz, daß *eine komplementäre Menge zu einer* (A)-*Menge selbst keine* (A)-*Menge zu sein braucht*; vielmehr sind in diesem Falle beide Mengen Borelsche. Herr Urysohn hat endlich *ein Beispiel einer im Einheitskreise regulären Potenzreihe gegeben, deren Randwertmenge keine Borelsche ist*; vielmehr beweist er, daß diese Menge stets eine (A)-Menge ist. In dem Begriffe der (A)-Mengen wird endlich ein den modernen Forderungen der Analysis vollständig angepaßter und in sich abgeschlossener Mengenvorrat gewonnen. Literatur: Souslins und Lusins Comptes-Rendus-Noten 1917, Lusin und Sierpiński: Bull. Ac. Polon. 1918, Journ. de Math. 1923; man vergl. auch außer meiner schon erwähnten C.-R.-Note zwei von mir verfaßte Artikel (Fundamenta Math. 5, Rec. Math. Moscou 31, 2) über die Komplementärmengen der (A)-Mengen.

lassen. *Gegeben sei ein System* $F_{i_1 i_2 \dots i_k}$ *von abgeschlossenen Mengen, wo* $k, i_1, i_2, \dots, i_k$ *alle Werte* $1, 2, \dots, n, \dots$ *durchlaufen. Die Menge, die als Vereinigungsmenge aller Mengen*

$$H_{i_1 i_2 \dots i_k \dots} = F_{i_1} \cdot F \quad \dots F_{i_1 i_2 \dots i_k} \dots$$

entsteht, ist dann eine (*A*)-*Menge.* Die ganze Theorie der (*A*)- und (*B*)-Mengen ist in topologischen bikompakten Räumen, in denen eine jede abgeschlossene Menge Durchschnitt von abzählbar vielen Gebieten ist, gültig. Es besteht insbesondere der Satz:

Satz VI. *In einem bikompakten Raume, in dem jedes* F *ein* G_δ *ist, enthält eine unabzählbare* (*A*)-*Menge* (*insbesondere* (*B*)-*Menge*) *stets eine perfekte Teilmenge und besitzt infolgedessen die Mächtigkeit* $\mathfrak{c}$.

Der Gedankengang des Beweises bleibt derselbe, wie er bereits im Falle des Euklidischen Raumes von mir im Jahre 1916 gegeben worden ist. Die Voraussetzung, daß ein jedes F ein G_δ ist, ist für die ganze Mächtigkeitstheorie sowie auch für die Theorie der (*A*)- und (*B*)-Mengen von grundlegender Bedeutung, wie man es an Gegenbeispielen sofort erkennt[14]).

Göttingen, den 3. Juli 1923.

P.S. Die Hauptergebnisse dieser Arbeit sind im März und Juni 1922 der Moskauer Mathematischen Gesellschaft und am 26. Juni 1923 der Göttinger Mathematischen Gesellschaft mitgeteilt worden.

[14]) Ich erwähne noch, daß der Mächtigkeitssatz für die Borelschen Mengen (unter gewöhnlichen Voraussetzungen) fast gleichzeitig mit mir von Herrn Hausdorff bewiesen worden ist (Math. Annalen 1916). Unter rein topologischen Voraussetzungen kommen meines Wissens Mächtigkeitsfragen erst heute in Betracht.

(Eingegangen am 1. 8. 1923.)

Über die Metrisation der kompakten topologischen Räume.

Von

Paul Urysohn † in Moskau.

1. Der Zweck dieser Arbeit ist zu beweisen:

Ein kompakter topologischer Raum[1]) *ist dann und nur dann metrisierbar, wenn er dem* II. *Abzählbarkeitsaxiom*[2]) *genügt*[3]); dabei nennen wir einen topologischen Raum *metrisierbar*, wenn zwischen seinen Punkten eine den gewöhnlichen Voraussetzungen genügende[4]) und zu denselben Limesbildungen wie das ursprüngliche Umgebungssystem führende Entfernung definiert werden kann.

Der zu beweisende Satz kann also auch folgendermaßen formuliert werden:

Das II. *Abzählbarkeitsaxiom ist eine notwendige und hinreichende Bedingung, damit ein kompakter topologischer Raum einem metrischen Raume*[5]) *homöomorph sei.*

Der Schwerpunkt ist der Beweis des Hinreichens (d. h. die *Metrisation*); auf den leicht zu führenden Beweis der Notwendigkeit gehe ich überhaupt nicht ein, da man ihn *implizite* schon im Hausdorffschen Buche finden kann[6]). Ebenso lasse ich beiseite die Erläuterung der Bedeutung, die dieser Satz für die Begründung der Topologie hat[7]).

---

[1]) Im Hausdorffschen Sinne: *Grundzüge der Mengenlehre*, Leipzig, Veit (1914), S. 213 und 230.

[2]) Hausdorff, l. c. S. 263 (Axiom (F)).

[3]) Allerdings muß das II. Abzählbarkeitsaxiom so formuliert werden, daß die aus endlich vielen Punkten bestehenden Räume dabei nicht, wie es bei Hausdorff der Fall ist, ausgeschaltet werden. Da jedoch solche Räume trivial sind, so genügt es, wenn wir im folgenden nur aus unendlich vielen Punkten bestehende Räume betrachten.

[4]) l. c. S. 211.

[5]) l. c. S. 211.

[6]) l. c. S. 274, X u. S. 273.

[7]) Man vergleiche hierzu eine Note, die ich vor kurzem in dem *Bull. de l'Acad. Polonaise* (1923) publiziert habe. — Während der Drucklegung dieser Arbeit ist eine

Bezeichnungen. Ich schließe mich im wesentlichen an die Hausdorffsche Terminologie an, benutze aber etwas andere Operationszeichen. Nämlich, die Vereinigungsmenge von A und B soll immer (auch wenn A und B nicht elementenfremd sind) durch $A+B$, der Durchschnitt dieser Mengen durch $A\times B$, ihre Differenz (d. h. die Menge der zu B nicht gehörenden Punkte von A; dabei wird nicht vorausgesetzt, daß B eine Teilmenge von A ist) durch $A-B$ bezeichnet werden. Das Inklusionszeichen $A\subset B$ bedeutet, daß A Teilmenge von B ist, wobei A und B auch zusammenfallen dürfen[8]).

2. Umgebungssystem. Es sei E ein dem II. Abzählbarkeitsaxiom genügender kompakter topologischer Raum, und

$$(1) \qquad V_1, V_2, \ldots, V_n, \ldots$$

ein ihn definierendes abzählbares Umgebungssystem[9]). Es ist zu bemerken, daß ein den Punkt x enthaltendes V_n nicht notwendig eine Umgebung *dieses* Punktes zu sein braucht. Wir führen also ein *neues* Umgebungssystem ein, wenn wir jedes V_n, das einen beliebigen Punkt x enthält, als Umgebung dieses Punktes bezeichnen; das Hausdorffsche Kriterium[10]) zeigt aber sofort, daß das neue System dem alten gleichwertig ist.

Wir werden also im folgenden immer voraussetzen, daß *jedes* V_n *eine Umgebung eines jeden in ihm enthaltenen Punktes ist.*

3. Überdeckungen. Da E kompakt ist, so gilt für ihn der Borelsche Überdeckungssatz[11]). Es gibt also gewiß solche *endliche* Systeme von Umgebungen

$$\Pi_i = \{V_i^{(1)}, V_i^{(2)}, \ldots, V_i^{(k_i)}\},$$

die den ganzen Raum überdecken; wir nennen sie *Überdeckungen*. Verschiedene Überdeckungen gibt es abzählbar viele; wir bezeichnen sie in irgendeiner Reihenfolge mit

$$(2) \qquad \Pi_1, \Pi_2, \ldots, \Pi_n, \ldots .$$

---

gemeinschaftliche Note von P. Alexandroff und mir in den Pariser *Comptes Rendus* (177, S. 1274) erschienen, in der unter Benutzung des (daselbst zitierten) Chittendenschen Satzes das Metrisationsproblem eine allgemeine (aber ziemlich komplizierte) Lösung findet. Übrigens erlaubt der Chittendensche Satz auch den zweiten Teil des Beweises des in dieser Arbeit behandelten Satzes erheblich zu vereinfachen: insbesondere hat mir Herr F. Hausdorff einen erstaunlich einfachen Beweis mündlich mitgeteilt. — In der soeben zitierten Note wird man auch weitere Literaturangaben finden.

[8]) $x\subset A$ besagt, daß der Punkt x zur Menge A gehört.

[9]) Da eine Menge ohne Umgebungen kein Raum ist, so ist es zweckmäßig zu sagen, daß das Umgebungssystem den Raum definiert. Verschiedene Umgebungssysteme in derselben Menge definieren im allgemeinen verschiedene Räume, wohl aber *denselben* Raum, wenn diese Systeme *gleichwertig* sind (Hausdorff, l. c., S. 260).

[10]) l. c. S. 261.

[11]) l. c. S. 272, VI.

Eine Überdeckung Π_i nennen wir *regulär*, wenn sie keine überflüssige Umgebungen enthält, d. h. wenn jedes $V_i^{(j)}$ mindestens einen Punkt (*Eigenpunkt*) enthält, der in keinem anderen $V_i^{(k)}$ liegt. Aus jedem nicht regulären Π_m kann man mittels sukzessiver Weglassung der überflüssigen Umgebungen ein reguläres $\Pi_{m'}$ gewinnen; die Beziehung von $\Pi_{m'}$ zu Π_m werden wir durch die Gleichung

$$\Pi_{m'} = J(\Pi_m)$$

bezeichnen. Wenn Π_m regulär ist, so soll $J(\Pi_m) = \Pi_m$ sein.

4. Hauptmenge. Es sei Π_n irgendeine reguläre Überdeckung. Wenn wir in jedem dazu gehörigen $V_n^{(j)}$ einen bestimmten Eigenpunkt wählen, so erhalten wir eine endliche Punktmenge; verfahren wir ebenso mit jeder regulären Überdeckung, so entsteht eine abzählbare Menge

$$a_1, a_2, \ldots, a_n, \ldots, \tag{3}$$

die wir ein für allemal fest gewählt denken, mit E_0 bezeichnen und *Hauptmenge* des Raumes E nennen.

E_0 ist im Raume E überall dicht[12]).

Es sei, in der Tat, V_n irgendeine Umgebung, und x irgendein in V_n liegender Punkt. Wir setzen dann $V_x = V_n$ und wählen für jeden von x verschiedenen Punkt y ein den Punkt x nicht enthaltendes V_y. Das so erhaltene unendliche Umgebungssystem $\{V_z\}_{z \subset E}$ überdeckt offenbar E; man kann folglich daraus ein Π_m, also auch ein reguläres Π_i herausgreifen. V_n wird aber dann notwendig zu diesem Π_i gehören, da es die einzige Umgebung des ursprünglichen Systems war, die den Punkt x enthielt. Zufolge der Konstruktion von E_0 enthält V_n — und n war willkürlich gewählt — mindestens einen Punkt von E_0, w. z. b. w.

Wenn wir jetzt irgendein reguläres Π_i

$$\Pi_i = \{V_i^{(1)}, V_i^{(2)}, \ldots, V_i^{(h_i)}\}$$

haben, so können wir für jedes seiner $V_i^{(j)}$ einen Eigenpunkt $a_i^{(j)}$ aus E_0 wählen. Wenn wir dabei jedesmal aus allen zulässigen Punkten denjenigen wählen, der in (3) den kleinsten Index hat, so werden wir sagen, daß

$$a_i^{(1)}, a_i^{(2)}, \ldots, a_i^{(h_i)}$$

die *Hauptpunkte* von Π_i sind.

5. Inklusion der Überdeckungen. Wir werden sagen, daß

$$\Pi_i = \{V_i^{(1)}, V_i^{(2)}, \ldots, V_i^{(h_i)}\}$$

in

$$\Pi_k = \{V_k^{(1)}, V_k^{(2)}, \ldots, V_k^{(h_k)}\}$$

---

[12]) l. c. S. 249.

enthalten ist — in Zeichen: $\Pi_i \subset \Pi_k$ —, wenn jedes $V_i^{(n)}$ in einem $V_k^{(m)}$ enthalten ist. Offenbar ist

$$\Pi_i \subset \Pi_i$$

und aus

$$\Pi_i \subset \Pi_k, \quad \Pi_k \subset \Pi_l$$

folgt

$$\Pi_i \subset \Pi_l.$$

Das Zusammenbestehen der beiden Inklusionen $\Pi_j \subset \Pi_i$ und $\Pi_j \subset \Pi_k$ werden wir kurz durch

$$(4) \qquad \Pi_j \subset \Pi_i \times \Pi_k$$

bezeichnen; und wir werden zeigen, daß es zu jedem Paare Π_i, Π_k ein der Inklusion (4) genügendes Π_j gibt.

In der Tat gehört jeder Punkt x von E zu einem $V_i^{(x)}$ aus Π_i und zu einem $V_k^{(x)}$ aus Π_k, also zu einem V_x, daß in $V_i^{(x)} \times V_k^{(x)}$ enthalten ist[13]). Wenn wir aber aus dem unendlichen Umgebungssystem $\{V_x\}_{x \subset E}$ ein Π_j

$$\Pi_j = \{V_{x_1}, V_{x_2}, \ldots, V_{x_n}\}$$

herausgreifen, so ist offenbar die Inklusion (4) erfüllt, da

$$V_{x_\nu} \subset V_i^{(x_\nu)}, \quad V_{x_\nu} \subset V_k^{(x_\nu)}$$

ist. (W. z. b. w.)

6. Kanonische Kette. Eine Folge von Überdeckungen

$$(5) \qquad \Pi_{n_1}, \Pi_{n_2}, \ldots, \Pi_{n_k}, \ldots$$

werden wir *kanonische Kette* nennen, wenn folgende vier Bedingungen erfüllt sind:

1. sämtliche Π_{n_k} sind regulär;
2. $\Pi_{n_k} \supset \Pi_{n_{k+1}}$ $(k = 1, 2, \ldots)$;
3. jeder Hauptpunkt von Π_{n_k} ist auch Hauptpunkt von $\Pi_{n_{k+1}}$;
4. jede Überdeckung Π_i enthält ein Π_{n_k}.

Wir beweisen jetzt, daß es kanonische Ketten wirklich gibt. Wir setzen hierzu $\Pi_{k_1} = J(\Pi_1)$ und verfahren weiter durch Induktion, wie folgt: Wenn $\Pi_{n_1}, \Pi_{n_2}, \ldots, \Pi_{n_{k-1}}$ schon definiert sind, so bezeichnen wir mit

$$(6) \qquad a_1, a_2, \ldots, a_{m_{k-1}}$$

ihre Hauptpunkte (d. h. die Hauptpunkte von $\Pi_{n_{k-1}}$); und es sei N_k der größte Index, den die Punkte (6) in der Folge (3) haben. Wir wählen dann eine der Inklusion

$$(7) \qquad \Pi_{(k)} \subset \Pi_{k-1} \times \Pi_{n_{k-1}}$$

---

[13]) Axiom (B) von Hausdorff.

genügende Überdeckung

$$\Pi_{(k)} = \{V_{(k)}^{(1)}, V_{(k)}^{(2)}, \ldots, V_{(k)}^{(t_k)}\}.$$

Wenn jetzt x irgendein Punkt von E ist, so seien

$$V_{(k)}^{(\mu_1)}, V_{(k)}^{(\mu_2)}, \ldots, V_{(k)}^{(\mu_r)}$$

alle diejenigen zu $\Pi_{(k)}$ gehörigen Umgebungen, die diesen Punkt enthalten; die Menge

$$(8) \qquad G_x = [V_{(k)}^{(\mu_1)} \times V_{(k)}^{(\mu_2)} \times \ldots \times V_{(k)}^{(\mu_r)}] - [\sum_{j=1}^{N_k} a_j - x]$$

ist offenbar ein den Punkt x enthaltendes Gebiet; also gibt es ein V_x, das in G_x enthalten ist. Das System $\{V_x\}_{x \subset E}$ überdeckt den Raum; dabei ist aber V_{a_i} $(i \leqq m_{k-1})$ die einzige Umgebung dieses Systems, die den Punkt a_i enthält. Wenn wir also aus dem System $\{V_x\}_{x \subset E}$ eine Überdeckung $\Pi_{[k]}$ herausgreifen, so werden alle V_{a_i} $(i \leqq m_{k-1})$ zu $\Pi_{[k]}$ gehören; dasselbe gilt auch für die reguläre Überdeckung $J(\Pi_{[k]})$, die wir mit Π_{n_k} bezeichnen wollen.

Aus der Konstruktion von $\Pi_{[k]}$ ist es sofort ersichtlich, daß $\Pi_{n_k} \subset \Pi_{[k]} \subset \Pi_{(k)}$, also, wegen (7), $\Pi_{n_k} \subset \Pi_{n_{k-1}}$ ist. Außerdem sind alle a_i $(i \leqq m_{k-1})$ Hauptpunkte von Π_{n_k}. In der Tat ist a_i ein Eigenpunkt der zu Π_{n_k} gehörenden Umgebung V_{a_i}; außerdem ist er in der Folge (3) enthalten und hat daselbst einen Index, der $\leqq N_k$ ist. Da aber zufolge (8) zwei verschiedene a_j $(j \leqq N_k)$ nicht zu einem V_x, also zu V_{a_i} gehören können, so ist a_i ein Hauptpunkt von Π_{n_k}.

Wenn wir jetzt die geschilderte Konstruktion weiter führen, so erhalten wir eine Folge (5), die, wie ersichtlich, den ersten drei Bedingungen genügt; daß aber auch die vierte erfüllt ist, sieht man sofort aus der für jedes k gültigen Inklusion

$$\Pi_{n_k} \subset \Pi_{(k)} \subset \Pi_{k-1}.$$

Damit ist unsere Behauptung bewiesen.

Es sei noch auf folgende zwei Tatsachen hingewiesen:

1) Sämtliche a_j, deren Indizes $\leqq N_k$ sind, sind Hauptpunkte von Π_{n_k}; also ist jeder Punkt von E_0 Hauptpunkt einer Überdeckung, die zur kanonischen Kette (5) gehört.

2) Jede Teilfolge der Folge (5) ist offenbar auch eine kanonische Kette, die die soeben genannte Eigenschaft besitzt.

7. Höhere Inklusionen. Wir verschärfen jetzt folgendermaßen den Begriff der Inklusion zweier Überdeckungen $\Pi_m = \{V_m^{(1)}, \ldots, V_m^{(h_m)}\}$ und $\Pi_n = \{V_n^{(1)}, \ldots, V_n^{(h_n)}\}$: wir werden erstens

$$\Pi_m \overset{2}{\subset} \Pi_n$$

schreiben, wenn es zu jeden zwei zueinander nicht fremden Umgebungen $V_m^{(i)}$ und $V_m^{(j)}$ von Π_m $(V_m^{(i)} \times V_m^{(j)} \neq 0)$ ein sie beide enthaltendes $V_n^{(l)}$ von Π_n gibt:

$$V_n^{(l)} \supset V_m^{(i)} + V_m^{(j)}.\text{[14]}$$

Wir werden zweitens

$$\Pi_m \overset{2}{\subset} \Pi_n$$

schreiben, wenn es zu jeden drei den beiden Bedingungen

$$V_m^{(i)} \times V_m^{(j)} \neq 0, \quad V_m^{(j)} \times V_m^{(k)} \neq 0$$

genügenden Umgebungen von Π_m ein $V_n^{(l)}$ von Π_n gibt, das sie sämtlich enthält:

$$V_n^{(l)} \supset V_m^{(i)} + V_m^{(j)} + V_m^{(k)}.$$

Es sei bemerkt, daß wir die Ungleichungen $i \neq j \neq k$ nicht vorausgesetzt haben; also folgt $\Pi_m \subset \Pi_n$ aus $\Pi_m \overset{2}{\subset} \Pi_n$ und $\Pi_m \overset{1}{\subset} \Pi_n$ aus $\Pi_m \overset{2}{\subset} \Pi_n$.

Es folgt weiter, wie leicht ersichtlich, sowie aus

$$\Pi_m \subset \Pi_n \overset{2}{\subset} \Pi_q,$$

wie auch aus

$$\Pi_m \overset{2}{\subset} \Pi_n \subset \Pi_q,$$

daß

$$\Pi_m \overset{2}{\subset} \Pi_q.$$

Endlich folgt aus den Inklusionen

$$\Pi_m \overset{2}{\subset} \Pi_n \overset{2}{\subset} \Pi_q,$$

daß

$$\Pi_m \overset{3}{\subset} \Pi_q;$$

in der Tat, wenn $V_m^{(i)} \times V_m^{(j)} \neq 0$ und $V_m^{(j)} \times V_m^{(k)} \neq 0$ sind, so gibt es, wegen $\Pi_m \overset{2}{\subset} \Pi_n$, ein $V_n^{(r)} \supset V_m^{(i)} + V_m^{(j)}$ und ein $V_n^{(s)} \supset V_m^{(j)} + V_m^{(k)}$, und es ist $V_n^{(r)} \times V_n^{(s)} \supset V_m^{(j)} \neq 0$; also gibt es, wegen $\Pi_n \overset{2}{\subset} \Pi_q$, ein

$$V_q^{(l)} \supset V_n^{(r)} + V_n^{(s)} \supset V_m^{(i)} + V_m^{(j)} + V_m^{(k)},$$

w. z. b. w.

8. Hilfssatz. Zu jedem Π_m gibt es ein $\Pi_n \overset{2}{\subset} \Pi_m$. Man kann dabei die Voraussetzung $\Pi_m \neq \Pi_n$ hinzufügen: es genügt in der Tat, den Hilfssatz auf ein von Π_m verschiedenes $\Pi_{m'} \subset \Pi_m$ anzuwenden.

Beweis. Wenn unsere Aussage nicht richtig wäre, so könnte man, insbesondere in jedem Π_{n_k} der kanonischen Kette (5), zwei zueinander nicht fremde Umgebungen $V_k^{\{1\}}$ und $V_k^{\{2\}}$ finden, deren Vereinigungsmenge $V_k^{\{1\}} + V_k^{\{2\}}$ in keinem $V_m^{(i)}$ von Π_m gänzlich enthalten ist. Es sei dann

[14]) Die Inklusion $\Pi \overset{2}{\subset} \Pi$ ist im allgemeinen nicht erfüllt.

y_k ein Punkt von $V_k^{(1)} \times V_k^{(2)}$, y irgendein (wegen der Kompaktheit von E sicher vorhandener) Häufungspunkt der Folge $\{y_k\}$; wir können aus $\{y_k\}$ eine gegen y konvergierende Teilfolge $\{y_{k_j}\}$ aussondern. Wir bezeichnen ferner mit V_m^y irgendeine zu Π_m gehörende Umgebung des Punktes y; alle y_{k_j}, deren Indizes j ein gewisses j_0 übertreffen, liegen dann in V_m^y. Unserer Voraussetzung gemäß muß mindestens eine der beiden Umgebungen $V_{k_j}^{(1)}$ und $V_{k_j}^{(2)}$ $(j > j_0)$ aus V_m^y herausragen; wir können immer die Bezeichnungen so wählen, daß dies mit der ersten geschieht:

$$V_{k_j}^{(1)} - V_m^y \neq 0 \qquad (j > j_0). \tag{9}$$

Wir definieren jetzt folgendermaßen ein System von Umgebungen $\{V_x\}_{x \subset E}$:

1. wenn $x \subset V_m^y$, so setzen wir $V_x = V_m^y$;
2. wenn $x \subset E - V_m^y$, so wählen wir ein $V_x \subset E - \{y + \sum_{j=j_0+1}^{\infty} y_{k_j}\}$.

Ein solches V_x existiert zufolge der Abgeschlossenheit von $y + \sum_{j=j_0+1}^{\infty} y_{k_j}$.

In diesem System $\{V_x\}$ ist folglich jedes V_x entweder mit V_m^y identisch oder zur Menge $\sum_{j>j_0} y_{k_j}$ fremd. Wenn wir also aus dem soeben definierten System eine reguläre Überdeckung $\tilde{\Pi} = \{V_{x_1}, V_{x_2}, \ldots, V_{x_s}\}$ herausgreifen, so ist eine der dazugehörigen Umgebungen mit V_m^y identisch, alle anderen aber zur Menge $\sum_{j=j_0+1}^{\infty} y_{k_j}$ fremd:

$$V_{x_1} = V_m^y, \qquad \sum_{t=2}^{s} V_{x_t} \times \sum_{j=j_0+1}^{\infty} y_{k_j} = 0.$$

Wir wählen endlich ein $\Pi_{n_\varkappa} \subset \tilde{\Pi}$ [15]) und ein den beiden Bedingungen

$$k_j \geqq \varkappa, \quad j > j_0$$

genügendes j; dann ist auch

$$\Pi_{n_{k_j}} \subset \tilde{\Pi}.$$

Infolgedessen muß $V_{k_j}^{(1)}$ in einem V_{x_t} enthalten sein; da aber $V_{k_j}^{(1)}$ den Punkt y_{k_j} $(j > j_0)$ enthält, so ist $t = 1$, also

$$V_{k_j}^{(1)} \subset V_{x_1} = V_m^y,$$

was mit (9) im Widerspruche steht. Damit ist unser Hilfssatz bewiesen.

9. Metrisierende Kette. Es sei $\Pi_{n_{k_1}}$ irgendeine zur kanonischen Kette (5) gehörende Überdeckung. Durch zweimalige Anwendung des so-

[15]) Siehe § 6, 4.

eben bewiesenen Hilfssatzes erhalten wir dann ein $\Pi_{k'} \overset{2}{\subset} \Pi_{n_{k_1}}$ und ein $\Pi_{k''} \overset{2}{\subset} \Pi_{k'}$; wenn außerdem $\Pi_{n_{k_2}}$ die erste Überdeckung aus (5) ist, die den beiden Bedingungen

$$k_2 > k_1, \quad \Pi_{n_{k_2}} \subset \Pi_{k''}$$

genügt, so ist (§ 7)

$$\Pi_{n_{k_2}} \overset{2}{\subset} \Pi_{k'} \overset{2}{\subset} \Pi_{n_{k_1}},$$

also

$$\Pi_{n_{k_2}} \overset{1}{\subset} \Pi_{n_{k_1}} \qquad (k_2 > k_1).$$

Wenn wir dieses Verfahren unendlich oft wiederholen, so erhalten wir eine Folge

$$(10^*) \qquad \Pi_{n_{k_1}}, \Pi_{n_{k_2}}, \ldots, \Pi_{n_{k_j}}, \ldots$$

von Überdeckungen, die erstens — zufolge der Schlußbemerkung des § 6 eine kanonische Kette bilden, zweitens aber nicht nur der allgemeinen Bedingung 2. für kanonische Ketten, sondern auch der schärferen Bedingung

$$2^*. \quad \Pi_{n_{k_j}} \overset{1}{\supset} \Pi_{n_{k_{j+1}}}$$

Genüge leisten.

Eine solche den Bedingungen 1., 2*., 3., 4. genügende Folge (10*) nennen wir *metrisierende Kette*.

Wir denken uns jetzt, daß die soeben gewonnene metrisierende Kette

$$(10) \qquad \Gamma_1, \Gamma_2, \ldots, \Gamma_n, \ldots$$

ein für allemal fest gewählt ist, und führen neue Bezeichnungen ein, indem wir

$$\left.\begin{array}{l} \Gamma_1 = \{V_1^1, V_2^1, \ldots, V_{h_1}^1\} \\ \Gamma_2 = \{V_1^2, V_2^2, \ldots, V_{h_2}^2\} \\ \ldots\ldots\ldots\ldots \\ \Gamma_n = \{V_1^n, V_2^n, \ldots, V_{h_n}^n\} \\ \ldots\ldots\ldots\ldots \end{array}\right\} (h_1 \leqq h_2 \leqq \ldots \leqq h_n \leqq \ldots \to \infty)$$

setzen, wobei außerdem die Bezeichnungen so gewählt sind, daß

$$V_i^n \supset V_i^{n+1} \supset \ldots \supset V_i^{n+\nu} \supset \ldots$$

und alle diese Umgebungen denselben Hauptpunkt a_i besitzen.

Die Hauptpunkte bezeichnen wir mit

$$\begin{array}{llll} a_1^1 &, a_2^1 &, \ldots, & a_{h_1}^1, \\ a_{h_1+1}^2 &, a_{h_1+2}^2 &, \ldots, & a_{h_2}^2, \\ \ldots & \ldots & \ldots & \ldots \\ a_{h_{n-1}+1}^n, & a_{h_{n-1}+2}^n, & \ldots, & a_{h_n}^n, \\ \ldots & \ldots & \ldots & \ldots \end{array}$$

wobei a_k^m der Hauptpunkt von $V_k^m, V_k^{m+1}, \ldots, V_k^{m+\mu}, \ldots$ ist; Hauptpunkte von Γ_n sind diejenigen a_k^m, bei denen $m \leqq n$ (oder $k \leqq h_n$) ist. Bei den a_k^m darf der obere Index auch fortgelassen werden; die Folge

$$(3^*) \qquad a_1, a_2, \ldots, a_k, \ldots$$

ist mit der ursprünglichen Folge (3) bis auf eine Umnumerierung identisch (§ 6, Schlußbemerkung):

$$\sum_{k=1}^{\infty} a_k = E_0.$$

10. Ordnungen. Wir nennen *Ordnung* einer Umgebung V_k^m bzw. eines Hauptpunktes a_k^m den betreffenden oberen Index m. Es ist zu bemerken, daß die Ordnung des Hauptpunktes einer Umgebung die Ordnung der Umgebung selbst nicht übertrifft; wenn a_k^m und $n \geqq m$ gegeben sind, so gibt es immer eine Umgebung n-ter Ordnung, deren Hauptpunkt a_k^m ist.

Es sei noch auf folgende Eigenschaft der metrisierenden Kette (10) hingewiesen:

Wenn V_k^n einen Hauptpunkt a_k^m von niedrigerer Ordnung besitzt ($m < n$; also ist V_k^{n-1} vorhanden), und wenn außerdem

$$V_k^n \times V_l^n \neq 0, \qquad V_l^n \times V_s^n \neq 0$$

ist, so ist

$$V_k^n + V_l^n + V_s^n \subset V_k^{n-1}.$$

In der Tat ist wegen der Bedingung 2*. (§ 9) $V_k^n + V_l^n + V_s^n$ in einem $V_\varkappa^{n-1}$ enthalten; da aber $V_\varkappa^{n-1} \supset V_k^n \supset a_k^m$ und a_k^m der Hauptpunkt von V_k^{n-1} ist, so muß $V_\varkappa^{n-1}$ mit V_k^{n-1} identisch sein. (W. z. b. w.)

11. Der Raum M_0. Es sei

$$M_0 = \{b_1, b_2, \ldots, b_k, \ldots\}$$

eine abzählbare Menge, die wir mit Hilfe der Relation $b_k \sim a_k$ eindeutig auf E_0 abbilden; wenn $a_k = a_k^m$ ist, so werden wir auch b_k^m statt b_k schreiben. Wir führen jetzt folgendermaßen *Entfernungen* $\varrho(b_k, b_l)$ ein[16]):

A_1. Wenn $V_k^1 \times V_l^1 = 0$ ist, so setzen wir $\varrho(b_k^1, b_l^1) = 1$;
wenn $V_k^1 \times V_l^1 \neq 0$ ist, so setzen wir $\varrho(b_k^1, b_l^1) = \frac{1}{2}$.

Damit erhalten wir für die h_1 Punkte b_k^1 erster Ordnung eine dem Dreiecksaxiom genügende Entfernung.

A_n. Wir nehmen jetzt an, daß eine dem Dreiecksaxiom genügende Entfernuug $\varrho(b_k, b_l)$ schon zwischen allen Punkten der Ordnung $< n$ (d. h. für $k, l \leqq h_{n-1}$) festgesetzt worden ist, und zwar in solcher Weise, daß die beiden Bedingungen

[16]) Da $\varrho(b_k, b_k) = 0$ zu setzen ist, so genügt es im folgenden, nur den Fall $k \neq l$ zu berücksichtigen.

(I_{n-1}): $\varrho(b_k, b_l)$ ist ein ganzes Vielfaches von $\frac{1}{2^{n-1}}$,

(II_{n-1}): $\varrho(b_k, b_l) = \frac{1}{2^{n-1}}$, wenn $V_k^{n-1} \times V_l^{n-1} \neq 0$

erfüllt sind ((I_1) und (II_1) waren, zufolge A_1, offenbar erfüllt).

Wir definieren jetzt $\varrho(b_k, b_l)$ für alle k und l, die $\leqq h_n$ sind; dabei soll diese Entfernung mit der alten übereinstimmen, wenn beide Punkte von einer Ordnung $< n$ sind, und es sollen sowohl das Dreiecksaxiom wie auch die den Bedingungen (I_{n-1}) und (II_{n-1}) analogen Bedingungen (I_n) und (II_n) erfüllt sein. Wir unterscheiden folgende drei Fälle:

α) Wenn $V_k^n \times V_l^n \neq 0$, so setzen wir $\varrho(b_k, b_l) = \frac{1}{2^n}$.

Diese Definition ist widerspruchsfrei, da mindestens einer der beiden Punkte von der Ordnung n ist. Wäre in der Tat die Ordnung von b_k kleiner als n, so wäre a_k Hauptpunkt einer Umgebung V_k^{n-1}, die (§ 10) $V_k^n + V_l^n$, also auch a_l enthalten würde; also wäre in diesem Falle a_l gewiß kein Hauptpunkt von Γ_{n-1}, d. h. kein Punkt von einer Ordnung $< n$.

β) Wenn $V_k^n \times V_l^n = 0$, aber $V_k^n \subset V_l^{n-1}$ ist, so setzen wir $\varrho(b_k, b_l) = \frac{1}{2^{n-1}}$.

Auch diese Definition ist widerspruchsfrei, da man sich — ebenso wie im vorigen Falle — sofort überzeugt, daß die Ordnung von b_k gleich n ist.

γ) Wir betrachten jetzt den Fall, wo die vorhergehenden Definitionen noch nicht genügen, $\varrho(b_k, b_l)$ festzulegen. Zwischen den verschiedenen die Punkte b_k und b_l verbindenden endlichen Punktketten $[b_k\, b_{s_1}\, b_{s_2} \ldots b_{s_i}\, b_l]$ gibt es dann gewiß auch solche — wir werden sie *metrische Ketten* nennen —, in denen für jedes benachbarte Punktepaar die Entfernung schon definiert ist.

Es sei in der Tat $V_{s_1}^{n-1}$ eine Umgebung von der Ordnung $n-1$, die V_k^n enthält ($V_k^n \supset a_k \sim b_k$)[17]), und $V_{s_2}^{n-1}$ eine solche, die $\supset V_l^n$; dann ist $[b_k\, b_{s_1}\, b_{s_2}\, b_l)$ eine metrische Kette.

Wir nennen *Länge* einer metrischen Kette die Summe der Längen ihrer Glieder, d. h. die Summe

$$\varrho(b_k, b_{s_1}) + \varrho(b_{s_1}, b_{s_2}) + \ldots + \varrho(b_{s_i}, b_l).$$

Ich behaupte jetzt, daß es zwischen b_k und b_l eine *Minimalkette*, d. h. eine metrische Kette kleinstmöglicher Länge gibt. In der Tat kann man mittels sukzessiver Weglassung der hintereinanderstehenden Wiederholungen eines und desselben Punktes und der geschlossenen Zyklen jede Kette durch eine nicht längere Kette ersetzen, die jeden Punkt höchstens

[17]) Wenn $k \leqq h_{n-1}$, d. h. wenn die Ordnung von $a_k < n$ ist, so ist $s_1 = k$.

einmal enthält. Da es aber solcher wiederholungsfreier Ketten offenbar nur endlich viele gibt, so ist unsere Behauptung bewiesen.

Wir setzen alsdann $\varrho(b_k, b_l)$ gleich der Länge einer Minimalkette.

Da die Bedingungen (I_n) und (II_n) offenbar erfüllt sind, so bleibt es nur zu zeigen, daß

$$\varrho(b_k, b_l) \leqq \varrho(b_k, b_t) + \varrho(b_t, b_l)$$

ist (Dreiecksaxiom). Da die Länge einer die Punkte b_k und b_t verbindenden Minimalkette $[b_k\, b_{s_1}\, b_{s_2} \ldots b_{s_i}\, b_t]$ unserer Definition gemäß jedenfalls nicht größer als $\varrho(b_k, b_t)$ sein kann[18]), ebenso die Länge der Minimalkette $[b_t\, b_{s_{i+1}}\, b_{s_{i+2}} \ldots b_{s_j}\, b_l] \leqq \varrho(b_t, b_l)$ ist, endlich aber die Länge einer b_k und b_l verbindenden Minimalkette $\leqq$ der Länge von $[b_k\, b_{s_1}\, b_{s_2} \ldots b_{s_i}\, b_t\, b_{s_{i+1}} \ldots b_{s_j}\, b_l]$, also $\leqq \varrho(b_k, b_t) + \varrho(b_t, b_l)$ ist, so genügt es, zu zeigen, daß

die Länge einer Minimalkette zwischen b_k und b_l nicht kleiner als $\varrho(b_k, b_l)$ sein kann.

Letztere Tatsache ist aber in den Fällen $A_n\gamma$, $A_n\alpha$ und $A_n\beta$ unmittelbar einleuchtend: im ersten zufolge der Definition von $\varrho(b_k, b_l)$ in diesem Falle, im zweiten infolgedessen, daß die Länge *jeder* Kette $\geqq \frac{1}{2^n}$ ist; im dritten Falle folgt sie endlich daraus, daß nur die eingliedrigen, auf Grund von $A_n\alpha$ — was hier offenbar nicht zutreffen kann — definierten Ketten eine Länge von $< \frac{1}{2^{n-1}}$ besitzen können.

Es bleibt also nur der Fall A_{n-1} übrig, d. h. derjenige, wo b_k und b_l beide eine Ordnung $< n$ haben. Wir betrachten in diesem Falle irgendeine metrische Kette $[b_k\, b_{s_1}\, b_{s_2} \ldots b_{s_i}\, b_l]$; wenn $b_k, b_{s_{\alpha_1}}, b_{s_{\alpha_2}}, \ldots, b_{s_{\alpha_\mu}}, b_l$ diejenigen Punkte dieser Kette sind, deren Ordnungen $< n$ sind, so ist zufolge dem, wie wir vorausgesetzt haben, für Punkte $< n$-ter Ordnung erfüllten Dreiecksaxiome

$$\varrho(b_k, b_l) \leqq \varrho(b_k, b_{s_{\alpha_1}}) + \varrho(b_{s_{\alpha_1}}, b_{s_{\alpha_2}}) + \cdots + \varrho(b_{s_{\alpha_\mu}}, b_l);$$

es genügt also zu zeigen, daß jedes $\varrho(b_{s_{\alpha_j}}, b_{s_{\alpha_{j+1}}})$ nicht größer als die Länge des betreffenden Abschnittes unserer Kette ist; d. h. wir haben nur solche Ketten zu betrachten, deren Endpunkte $< n$-ter, sämtliche Mittelpunkte aber von n-ter Ordnung sind.

Es sei also $[b_k\, b_{s_1}\, b_{s_2} \ldots b_{s_i}\, b_l]$ eine solche Kette:

$$k, l \leqq h_{n-1}, \quad s_j > h_{n-1} \qquad (j = 1, 2, \ldots, i);$$

[18]) In den Fällen A_{n-1} (d. i. $k, t \leqq h_{n-1}$), $A_n\alpha$ und $A_n\beta$ ist diese Kette eingliedrig.

wenn wir ihre Länge durch $\frac{r}{2^n}$ bezeichnen[19]), so haben wir zu zeigen, daß $\varrho(b_k, b_l) \leqq \frac{r}{2^n}$ ist.

Beweis. Zufolge der Definition der metrischen Ketten ist die Entfernung zwischen b_{s_j} und $b_{s_{j+1}}$ notwendigerweise auf Grund von $A_n\,\alpha$ eingeführt; also ist

$$\varrho(b_{s_j}, b_{s_{j+1}}) = \frac{1}{2^n}, \quad V^n_{s_j} \times V^n_{s_{j+1}} \neq 0 \quad (j = 1, 2, \ldots, i-1).$$

Für $\varrho(b_k, b_{s_1})$ sind hingegen zwei Möglichkeiten vorhanden: entweder sind wir auch hier im Falle $A_n\,\alpha$, — und dann ist $\varrho(b_k, b_{s_1}) = \frac{1}{2^n}$ und $V^n_k \times V^n_{s_1} \neq 0$; oder im Falle $A_n\,\beta$, was $\varrho(b_k, b_{s_1}) = \frac{2}{2^n}$ und $V^{n-1}_k \supset V^n_{s_1}$ zur Folge hat. Dasselbe gilt auch vom Punktepaare b_{s_i}, b_l.

Wir können voraussetzen, daß $r \geqq 3$ ist: der Fall einer eingliedrigen Kette $[b_k b_l]$ ist nämlich trivial, und in allen anderen Fällen ist die genannte Voraussetzung erfüllt. Es könnte in der Tat $r < 3$ nur noch im Falle einer zweigliedrigen Kette $[b_k b_{s_1} b_l]$ sein, wenn außerdem $\varrho(b_k, b_{s_1}) = \varrho(b_{s_1}, b_l) = \frac{1}{2^n}$ wäre; dies ist aber unmöglich, da die, wie wir soeben gesehen haben, in diesem Falle erfüllten Relationen $V^n_k \times V^n_{s_1} \neq 0$ und $V^n_{s_1} \times V^n_l \neq 0$ die Existenz einer $V^{n-1}_\lambda \supset V^n_k + V^n_{s_1} + V^n_l$, also einer Umgebung $(n-1)$-ter Ordnung, die zwei Hauptpunkte, a_k und a_l, von Γ_{n-1} enthält, zur Folge hätten.

Wir führen jetzt folgende Bezeichnungen ein: Wenn die Länge des Abschnittes $[b_k b_{s_1} \ldots b_{s_j}]$ unserer Kette $= \frac{\alpha}{2^n}$ ist, so setzen wir $b_{s_j} = b_{[\alpha]}$, $a_{s_j} = a_{[\alpha]}$ und $V^n_{s_j} = V^n_{[\alpha]}$; also ist $b_k = b_{[0]}$, $b_{s_1} = b_{[1]}$ oder $b_{[2]}$, $b_{s_i} = b_{[r-2]}$ oder $b_{[r-1]}$, $b_l = b_{[r]}$.

Ich behaupte, daß $V^{n-1}_k \supset a_{[2]}$: wenn $b_{[1]}$ nicht existiert, also $b_{[2]} = b_{s_1}$, $\varrho(b_k, b_{s_1}) = \frac{2}{2^n}$ ist, so ist $V^{n-1}_k \supset V^n_{s_1} \supset a_{s_1} = a_{[2]}$; und im Falle der Existenz von $b_{[1]}$ ist $b_{s_1} = b_{[1]}$, $b_{s_2} = b_{[2]}$, $\varrho(b_k, b_{s_1}) = \varrho(b_{s_1}, b_{s_2}) = \frac{1}{2^n}$, also $V^n_k \times V^n_{s_1} \neq 0$, $V^n_{s_1} \times V^n_{s_2} \neq 0$, was (§ 10) die Inklusion

$$V^{n-1}_k \supset V^n_k + V^n_{s_1} + V^n_{s_2} \supset a_{s_2} = a_{[2]}$$

zur Folge hat.

---

[19]) r ist zufolge (I_n) eine ganze Zahl.

Ebenso wird die Inklusion $V_l^{n-1} \supset a_{[r-2]}$ bewiesen[19a]).

Es sei jetzt σ irgendeine den Ungleichungen

$$0 < \sigma < \tau = E\left(\frac{r-3}{2}\right)^{20)}$$

genügende ganze Zahl. Dann ist, wie wir schon gesehen haben,

$$V_{[2\sigma]}^n \times V_{[2\sigma+1]}^n \neq 0, \qquad V_{[2\sigma+1]}^n \times V_{[2\sigma+2]}^n \neq 0;$$

es folgt also aus $\Gamma_{n-1} \overset{3}{\supset} \Gamma_n$, daß es eine Umgebung $(n-1)$-ter Ordnung $V_{\{\sigma\}}^{n-1}$ gibt, die $V_{[2\sigma]}^n + V_{[2\sigma+1]}^n + V_{[2\sigma+2]}^n$ enthält; folglich ist

$$V_{\{\sigma\}}^{n-1} \supset a_{[2\sigma]} + a_{[2\sigma+2]}.$$

Da $(r-2) - 2\tau \leqq 2$ ist, so können wir in derselben Weise auf die Existenz einer der Inklusion

$$V_{\{\tau\}}^{n-1} \supset a_{[2\tau]} + a_{[r-2]}$$

genügenden Umgebung $(n-1)$-ter Ordnung $V_{\{\tau\}}^{n-1}$ schließen.

Wir betrachten jetzt die Punktkette

$$(11) \qquad [b_k\, b_{\{1\}}\, b_{\{2\}} \ldots b_{\{\tau\}}\, b_l],$$

wo $b_{\{\sigma\}} \sim a_{\{\sigma\}}$ und $a_{\{\sigma\}}$ der Hauptpunkt von $V_{\{\sigma\}}^{n-1}$ ist. Sie besteht aus lauter Punkten $<n$-ter Ordnung und besitzt Glieder, deren Längen sämtlich $\leqq \frac{1}{2^{n-1}}$ sind[21]), wie es aus den Inklusionen

$$\begin{aligned} &V_k^{n-1} \times V_{\{1\}}^{n-1} \supset a_{[2]}, \\ &V_{\{\sigma\}}^{n-1} \times V_{\{\sigma+1\}}^{n-1} \supset a_{[2\sigma+2]} \qquad (\sigma = 1, 2, \ldots, \tau - 1), \\ &V_{\{\tau\}}^{n-1} \times V_l^{n-1} \supset a_{[r-2]} \end{aligned}$$

ersichtlich ist; da anderseits für Punkte $<n$-ter Ordnung das Dreiecksaxiom erfüllt ist, so kommt

---

[19a]) Im Falle $r = 3$ kann es vorkommen, daß einer der beiden Punkte $a_{[2]}$ und $a_{[r-2]} = a_{[1]}$ nicht existiert; einer von ihnen — es sei $a_{[\beta]}$ — ist aber sicher vorhanden, und es ist: $V_k^{n-1} \supset a_{[\beta]}$, $V_l^{n-1} \supset a_{[\beta]}$, also $V_k^{n-1} \times V_l^{n-1} \neq 0$, $\varrho(b_k, b_l) = \frac{1}{2^{n-1}} < \frac{r}{2^n}$, womit der Fall $r = 3$ vollständig erledigt ist und im folgenden nicht mehr berücksichtigt zu sein braucht.

[20]) $E(x)$ bedeutet die größte ganze Zahl, die $\leqq x$ ist; also ist $2\tau = r - 3$ oder $r - 4$.

[21]) Das Ungleichheitszeichen ist darum zu setzen, weil zwei konsekutive Punkte der Kette (11) auch zusammenfallen können.

$$\varrho(b_k, b_l) \leqq \varrho(b_k, b_{\{1\}}) + \sum_{\sigma=1}^{r-1} \varrho(b_{\{\sigma\}}, b_{\{\sigma+1\}}) + \varrho(b_{\{r\}}, b_l)$$

$$\leqq \frac{r+1}{2^{n-1}} = \frac{E\left(\frac{r-3}{2}\right)+1}{2^{n-1}} \leqq \frac{r-1}{2^n} < \frac{r}{2^n}, \qquad \text{w. z. b. w.}$$

Wir haben also ein Induktionsverfahren erhalten, mit Hilfe dessen wir für alle Punktepaare aus M_0 eine dem Dreiecksaxiom genügende Entfernung definieren; M_0 wird damit in einen *metrischen Raum* verwandelt.

12. Relationen zwischen E_0 und M_0. Es sei b_k^m ein Punkt von der Ordnung $m > n$, und $V_k^m \supset a_k^m \sim b_k^m$; wir können dann eine Reihe von Umgebungen $V_{k_1}^{m-1}, V_{k_2}^{m-2}, \ldots, V_{k_{m-n}}^n$ derart finden, daß

$$V_k^m \subset V_{k_1}^{m-1} \subset V_{k_2}^{m-2} \subset \ldots \subset V_{k_{m-n}}^n \tag{12}$$

Wenn $a_{k_1}, a_{k_2}, \ldots, a_{k_{m-n}}$ die Hauptpunkte dieser Umgebungen sind, so bezeichnen wir den Punkt $b_{k_{m-n}}$ (dessen Ordnung $\leqq n$ ist) mit $c_n(b_k^m)$; $c_n(b_k)$ ist allerdings im allgemeinen durch b_k und n nicht eindeutig bestimmt, was aber für uns belanglos ist. Wenn $m \leqq n$ ist, so setzen wir $c_n(b_k^m) = b_k^m$. Wir beweisen jetzt, daß

$$\varrho(b_k, c_n(b_k)) < \frac{1}{2^{n-1}} \tag{13}$$

ist. Es genügt, den Fall $m > n$ zu betrachten; es folgt dann aus (12) und der Entfernungsdefinition des § 11, daß

$$\varrho(b_k, b_{k_1}) \leqq \frac{1}{2^{m-1}}, \quad \varrho(b_{k_1}, b_{k_2}) \leqq \frac{1}{2^{m-2}}, \ldots, \varrho(b_{k_{m-n-1}}, b_{k_{m-n}}) \leqq \frac{1}{2^n},$$

also

$$\varrho(b_k, b_{k_{m-n}}) \leqq \frac{1}{2^n} + \frac{1}{2^{n+1}} + \ldots + \frac{1}{2^{m-1}} < \frac{1}{2^{n-1}}. \qquad \text{(W. z. b. w.)}$$

Es sei noch auf folgende, aus unserer Konstruktion unmittelbar ersichtliche Tatsache hingewiesen:

$$\text{aus } b_{k'} = c_n(b_k) \text{ folgt } V_{k'}^n \supset a_k.$$

Satz N. *Wenn a_k und a_l zu derselben Umgebung von der Ordnung n gehören, so ist $\varrho(b_k, b_l) < \frac{1}{2^{n-3}}$.*

Beweis. Es sei V_λ^n diejenige Umgebung, die $a_k + a_l$ enthält, und es sei ferner $b_{k'} = c_n(b_k)$, $b_{l'} = c_n(b_l)$. Dann ist

$$V_{k'}^n \times V_\lambda^n \supset a_k \neq 0, \qquad V_\lambda^n \times V_{l'}^n \supset a_l \neq 0,$$

also existiert ein $V_\mu^{n-1} \supset V_{k'}^n + V_\lambda^n + V_{l'}^n$. Der dem Hauptpunkte a_μ von

V_μ^{n-1} entsprechende Punkt b_μ ist folglich gleichzeitig ein $c_{n-1}(b_k)$ und ein $c_{n-1}(b_l)$, was wegen (13)

$$\varrho(b_k, b_l) \leqq \varrho(b_k, b_\mu) + \varrho(b_\mu, b_l) < 2 \cdot \frac{1}{2^{n-2}} = \frac{1}{2^{n-3}},$$

d. h. unsere Behauptung zur Folge hat.

Hilfssatz 1_n. Wenn b_k und b_l zwei Punkte von der Ordnung $\leqq n$ sind ($k, l \leqq h_n$), deren Entfernung $\varrho(b_k, b_l) \leqq \frac{2}{2^n}$ ist, so gibt es eine Umgebung von der Ordnung $n-2$, die die Punkte a_k und a_l gleichzeitig enthält.

Beweis. $\varrho(b_k, b_l)$ ist entweder auf Grund von $A_{n-1}\alpha$, oder von $A_n\alpha$, oder von $A_n\beta$, oder endlich von $A_n\gamma$ festgelegt. Im ersten Falle ist $V_k^{n-1} \times V_l^{n-1} \neq 0$, also existiert ein $V_\lambda^{n-2} \supset V_k^{n-1} + V_l^{n-1} \supset a_k + a_l$. Im zweiten existiert sogar ein $V_\mu^{n-1} \supset a_k + a_l$, also ein $V_\lambda^{n-2} \supset V_\mu^{n-1} \supset a_k + a_l$. Im dritten ist entweder $V_k^n \subset V_l^{n-1}$, also $V_l^{n-1} \supset a_k + a_l$, oder $V_l^n \subset V_k^{n-1}$, also $V_k^{n-1} \supset a_k + a_l$. Im vierten kann die zugehörige Minimalkette nur zweigliedrig sein, wobei, wenn $[b_k b_s b_l]$ eine solche Kette ist, sowohl $\varrho(b_k, b_s)$, als auch $\varrho(b_s, b_l)$ dem Falle $A_n\alpha$ entsprechen; es existieren also ein $V_\mu^{n-1} \supset a_k + a_s$ und ein $V_\tau^{n-1} \supset a_s + a_l$, also, wegen $V_\mu^{n-1} \times V_\tau^{n-1} \supset a_s \neq 0$, ein $V_\lambda^{n-2} \supset V_\mu^{n-1} + V_\tau^{n-1} \supset a_k + a_l$.

Hilfssatz 2_n. Wenn $k, l \leqq h_n$ und $\varrho(b_k, b_l) < \frac{5}{2^n}$ ist, so gibt es ein V_λ^{n-3}, das $a_k + a_l$ enthält.

Beweis. Da b_k und b_l beide von einer Ordnung $\leqq n$ sind, so ist $\varrho(b_k, b_l)$ ein ganzes Vielfaches von $\frac{1}{2^n}$, also $\leqq \frac{4}{2^n}$. Wir können voraussetzen, daß 1. $\varrho(b_k, b_l) \geqq \frac{3}{2^n}$, 2. die Ordnung wenigstens eines der beiden Punkte $= n$ ist: wäre nämlich eine dieser beiden Voraussetzungen nicht erfüllt, so wäre unsere Behauptung eine Folge des vorigen Hilfssatzes (in der Form 1_n bzw. 1_{n-1}). $\varrho(b_k, b_l)$ ist also auf Grund von $A_n\gamma$ definiert und die entsprechende Minimalkette ist höchstens viergliedrig. Wenn die Kette viergliedrig ist, so haben alle ihre Glieder die Länge $\frac{1}{2^n}$; jedenfalls ist aber die Länge jedes Gliedes $\leqq \frac{2}{2^n}$. Wir sehen also, daß man immer zwei solche (nicht unbedingt voneinander verschiedene) Punkte b_{s_1} und b_{s_2} ($s_1, s_2 \leqq h_n$) wählen kann, daß

$$\varrho(b_k, b_{s_1}) \leqq \frac{2}{2^n}, \quad \varrho(b_{s_1}, b_{s_2}) \leqq \frac{2}{2^n}, \quad \varrho(b_{s_2}, b_l) \leqq \frac{2}{2^n}$$

ist. Daraus folgt aber (Hilfssatz 1_n) die Existenz von $V_\mu^{n-2} \supset a_k + a_{s_1}$, von $V_\sigma^{n-2} \supset a_{s_1} + a_{s_2}$ und von $V_\tau^{n-2} \supset a_{s_2} + a_l$, also wegen $V_\mu^{n-2} \times V_\sigma^{n-2} \supset a_{s_1}$,

$V_\sigma^{n-2} \times V_\tau^{n-2} \supset a_{s_2}$, die Existenz eines $V_\lambda^{n-3} \supset V_\mu^{n-2} + V_\sigma^{n-2} + V_\tau^{n-2} \supset a_k + a_l$, w. z. b. w.

Satz T. *Wenn* $\varrho(b_k, b_l) < \frac{1}{2^n}$ *ist, so gibt es eine Umgebung von der Ordnung* $n-4$, *die beide Punkte* a_k *und* a_l *gleichzeitig enthält.*

Beweis. Wir benutzen wieder die am Anfange dieses Paragraphen angegebene Konstruktion: es sei nämlich $b_{k'} = c_n(b_k)$, $b_{l'} = c_n(b_l)$. Dann ist, wie wir gesehen haben,

$$V_{k'}^n \supset a_k + a_{k'}\ ^{22}), \qquad V_{l'}^n \supset a_l + a_{l'},$$

$$\varrho(b_k, b_{k'}) < \frac{2}{2^n}, \qquad \varrho(b_l + b_{l'}) < \frac{2}{2^n},$$

also

$$\varrho(b_{k'}, b_{l'}) \leqq \varrho(b_{k'}, b_k) + \varrho(b_k, b_l) + \varrho(b_l, b_{l'}) < \frac{5}{2^n}.$$

Da anderseits die Ordnung von $b_{k'}$ und $b_{l'} \leqq n$ ist, so folgt (Hilfssatz 2_n) aus dieser Ungleichung die Existenz eines $V_\sigma^{n-3} \supset a_{k'} + a_{l'}$. Wenn wir jetzt ein $V_\mu^{n-3} \supset V_{k'}^n$ und ein $V_\tau^{n-3} \supset V_{l'}^n$ wählen, so ist

$$V_\mu^{n-3} \times V_\sigma^{n-3} \supset V_{k'}^n \times V_\sigma^{n-3} \supset a_{k'}, \qquad V_\sigma^{n-3} \times V_\tau^{n-3} \supset a_{l'};$$

es gibt also ein $V_\lambda^{n-4} \supset V_\mu^{n-3} + V_\sigma^{n-3} + V_\tau^{n-3} \supset a_k + a_l$, w. z. b. w.

13. Fundamentalfolgen[23]).

Satz A. *Wenn die Folge* $\{a_{i_1}, a_{i_2}, \ldots, a_{i_\nu}, \ldots\}$ *in* E *konvergiert, so ist* $\{b_{i_1}, b_{i_2}, \ldots, b_{i_\nu}, \ldots\}$ *eine Fundamentalfolge.*

Es sei in der Tat x der Limes von $\{a_{i_\nu}\}$ und $n \geqq 3$ eine ganze Zahl. Wenn V_λ^n eine den Punkt x enthaltende Umgebung von der Ordnung n ist, so enthält sie alle a_{i_ν}, bei denen ν größer als ein gewisses ν_n ist; folglich ist (Satz N) $\varrho(b_{i_{\nu'}}, b_{i_{\nu''}}) < \frac{1}{2^{n-3}}$ für alle $\nu', \nu'' > \nu_n$, w. z. b. w.

Zusatz. Wenn $\{a_{i_\nu}\}$ und $\{a_{j_\nu}\}$ gegen denselben Punkt konvergieren, so sind $\{b_{i_\nu}\}$ und $\{b_{j_\nu}\}$ *konfinal*, d. h. $\varrho(b_{i_\nu}, b_{j_\nu})_{\nu\to\infty} \to 0$.

Die Folge $\{a_{i_1}, a_{j_1}, a_{i_2}, a_{j_2}, \ldots, a_{i_\nu}, a_{j_\nu}, \ldots\}$ ist in der Tat konvergent; die Anwendung des Satzes A liefert also sofort das gewünschte Ergebnis.

Hilfsatz. Wenn x und y zwei verschiedene Punkte von E sind, so gibt es ein Γ_n, welches keine einzige der Inklusion

$$x + y \subset V_\lambda^n$$

genügende Umgebung V_λ^n enthält.

---

[22]) $a_{k'}$ ist der Hauptpunkt von $V_{k'}^n$.

[23]) Hausdorff, l. c. S. 314.

Es seien V_x und V_y zwei zueinander fremde Umgebungen; zu jedem von x und y verschiedenen Punkte von E wählen wir dann eine in $E-(x+y)$ enthaltene Umgebung. Wenn wir aus dem in dieser Weise erhaltenen Umgebungssystem $\{V_z\}_{z\subset E}$ eine Überdeckung $\tilde{\Pi}$ herausgreifen und ein $\Gamma_n \subset \tilde{\Pi}$ wählen, so besitzt ersichtlich Γ_n die verlangte Eigenschaft.

Satz B. *Wenn* $\{b_{i_\nu}\}$ *eine Fundamentalfolge ist, so ist* $\{a_{i_\nu}\}$ *im Raume E konvergent.*

Zufolge der Kompaktheit von E genügt es, zu zeigen, daß $\{a_{i_\nu}\}$ einen einzigen Häufungspunkt hat[24]). Wir setzen hingegen voraus, daß x und y zwei verschiedene Häufungspunkte von $\{a_{i_\nu}\}$ sind, und wählen alsdann ein Γ_n, welches dem vorigen Hilfssatze entspricht; es gibt also kein V_λ^n, das x und y gleichzeitig enthält. Wir wählen dann ein solches ν_n, daß für alle $\nu', \nu'' > \nu_n$ die Entfernung $\varrho(b_{i_{\nu'}}, b_{i_{\nu''}}) < \frac{1}{2^{n+5}}$ ist, und zwei Umgebungen von der Ordnung $n+1$, nämlich $V_\xi^{n+1} \supset x$ und $V_\eta^{n+1} \supset y$. Es gibt, unserer Voraussetzung gemäß, Punkte a_{i_ν} mit beliebig hohen Indizes ν, die in V_ξ^{n+1} bzw. in V_η^{n+1} enthalten sind; es seien also ν' und ν'' zwei ganze Zahlen, die den Bedingungen

$$\nu', \nu'' > \nu_n, \qquad a_{i_{\nu'}} \subset V_\xi^{n+1}, \qquad a_{i_{\nu''}} \subset V_\eta^{n+1}$$

genügen; daraus folgt die Ungleichung $\varrho(b_{i_{\nu'}}, b_{i_{\nu''}}) < \frac{1}{2^{n+5}}$, also (Satz T) die Existenz eines $V_\mu^{n+1} \supset a_{i_{\nu'}} + a_{i_{\nu''}}$. Folglich ist

$$V_\xi^{n+1} \times V_\mu^{n+1} \supset a_{i_{\nu'}}, \qquad V_\mu^{n+1} \times V_\eta^{n+1} \supset a_{i_{\nu''}},$$

was auf die Existenz eines

$$V_\lambda^n \supset V_\xi^{n+1} + V_\mu^{n+1} + V_\eta^{n+1} \supset x + y,$$

also auf einen Widerspruch führt. Damit ist Satz B bewiesen.

Zusatz. Wenn $\{b_{i_\nu}\}$ und $\{b_{j_\nu}\}$ konfinale Fundamentalfolgen sind, so konvergieren $\{a_{i_\nu}\}$ und $\{a_{j_\nu}\}$ zu demselben Punkte.

Es genügt, den Satz B auf die Fundamentalfolge $\{b_{i_1}, b_{j_1}, \ldots, b_{i_\nu}, b_{j_\nu}, \ldots\}$ anzuwenden.

14. Der Raum M. Wir können M_0 als eine überall dichte Teilmenge eines *vollständigen metrischen Raumes M* betrachten[25]).

Es sei x irgendein Punkt von E, und $\{a_{i_\nu}\}$ eine gegen ihn konvergierende Folge[26]). Die Folge $\{b_{i_\nu}\}$ ist dann, wie wir es im vorigen Para-

[24]) l. c. S. 234.

[25]) Hausdorff, l. c. S. 315.

[26]) Die a_{i_ν} sind Punkte der in E überall dichten Menge E_0.

graphen gesehen haben, eine Fundamentalfolge; sie konvergiert also gegen einen Punkt y von M; und es folgt aus dem Zusatze zum Satze A, daß y eindeutig durch x bestimmt ist. Anderseits folgt aus der Dichtigkeit von M_0 in M und aus dem Satze B, daß *jeder* Punkt y von M auf diese Weise gewonnen werden kann. Der Zusatz zum letzteren Satze lehrt uns endlich, daß zweien verschiedenen Punkten x_1 und x_2 von E zwei verschiedene Punkte y_1 und y_2 von M entsprechen.

Wir haben also eine eineindeutige Abbildung von E auf M erhalten; diese Abbildung ist eine Erweiterung der ursprünglichen Abbildung von E_0 auf M_0, wie man es leicht aus der Betrachtung von Folgen $\{a_{i_1}, a_{i_2}, \ldots, a_{i_\nu}, \ldots\}$ $(a_i \subset E_0)$ schließen kann.

Ich behaupte endlich, daß diese Abbildung nebst ihrer Umkehrung *stetig* ist. Wegen der Kompaktheit von E genügt es, zu zeigen, daß M stetiges Bild von E ist[27]); dieses folgt aber unmittelbar aus dem folgenden Satze:

Es seien x_1 und x_2 zwei Punkte von E, die in einer und derselben Umgebung n-ter Ordnung V_λ^n liegen; wenn y_1 und y_2 die ihnen entsprechenden Punkte von M sind, so ist $\varrho(y_1, y_2) \leqq \frac{1}{2^{n-3}}$.

Es seien $\{a_{i_\nu}\} \to x_1$ und $\{a_{j_\nu}\} \to x_2$ zwei Folgen zu E_0 gehörender Punkte; dann ist unserer Konstruktion gemäß $\{b_{i_\nu}\} \to y_1$ und $\{b_{j_\nu}\} \to y_2$. Wir können also, da $x_1 + x_2 \subset V_\lambda^n$ ist, zu jedem $\varepsilon > 0$ ein solches ν wählen, daß

$$(14) \qquad a_{i_\nu} + a_{j_\nu} \subset V_\lambda^n,$$

$$\varrho(y_1, b_{i_\nu}) < \frac{\varepsilon}{2}, \qquad \varrho(y_2, b_{j_\nu}) < \frac{\varepsilon}{2}$$

ist. Aus (14) folgt aber (Satz N, § 12), daß

$$\varrho(b_{i_\nu}, b_{j_\nu}) < \frac{1}{2^{n-3}},$$

also

$$\varrho(y_1, y_2) \leqq \varrho(y_1, b_{i_\nu}) + \varrho(b_{i_\nu}, b_{j_\nu}) + \varrho(b_{j_\nu}, y_2) < \frac{1}{2^{n-3}} + \varepsilon.$$

Da dies für jedes ε gilt, so ist unsere Behauptung bewiesen.

Wir haben also bewiesen, daß es eine eineindeutige und beiderseits stetige Abbildung von E auf M gibt, d. h. daß *jeder dem*

---

[27]) l. c. S. 365, VIII.

II. *Abzählbarkeitsaxiom genügende kompakte topologische Raum einem metrischen Raume homöomorph ist.*

Damit ist unser Ziel erreicht.

Zum Schlusse sei noch bemerkt, daß Herr Paul Alexandroff viele interessante Anwendungen des soeben bewiesenen Satzes gefunden hat; insbesondere hat er, auf diesen Satz fußend, das Metrisationsproblem für die *im Kleinen kompakten* Räume gelöst[28]).

Göttingen, den 15. VII. 1923.

---

[28]) Diese Arbeit erscheint gleichzeitig in diesen Annalen.

(Eingegangen am 1. 8. 1923.)

Über die Metrisation der im Kleinen kompakten topologischen Räume[1]).

Von

Paul Alexandroff in Moskau.

1. Definitionen und Vorbemerkungen.

In einer vor kurzem erschienenen Arbeit habe ich die bikompakten topologischen Räume insbesondere in bezug auf ihre Struktureigenschaften im Kleinen untersucht. Es ist aber ohne weiteres klar, daß für die Gültigkeit dieser Eigenschaften die Bedingung der Bikompaktheit des gesamten Raumes überflüssig ist; es genügt vielmehr, die letzte Eigenschaft von gewissen, bestimmte Umgebungen sämtlicher Raumpunkte enthaltenden abgeschlossenen Mengen zu verlangen. In dieser Weise gelangen wir naturgemäß zu dem Begriffe der im Kleinen bikompakten (resp. kompakten) topologischen Räume, indem wir folgende grundlegende Definition aufstellen:

A. *Ein topologischer Raum $\mathfrak{R}$ heißt* ***im Punkte*** *ξ bikompakt* (bzw. *kompakt*), *falls eine gewisse Umgebung $U(\xi)$ dieses Punktes existiert, so daß die abgeschlossene Menge $\overline{U}(\xi)$, als Relativraum betrachtet, bikompakt* (bzw. *kompakt*) *ist*. ($\overline{U}(\xi)$ bedeutet hier, wie immer, die kleinste die Menge $U(\xi)$ enthaltende, in $\mathfrak{R}$ abgeschlossene Menge, d. h.

$$\overline{U}(\xi) = U(\xi) + [U(\xi)]',$$

wo $[U(\xi)]'$ die Menge aller Häufungspunkte von $U(\xi)$ ist.)

B. *Ein in jedem seiner Punkte bikompakter* (bzw. *kompakter*) *topologischer Raum heißt* ***im Kleinen bikompakt*** (bzw. ***kompakt***).

---

[1]) Außer den für die ganze Theorie grundlegenden Hausdorffschen Definitionen (Grundzüge der Mengenlehre, Leipzig (1914), Kap. VII, VIII), werden die Bezeichnungen und Ergebnisse der folgenden voranstehenden Arbeiten als bekannt vorausgesetzt:

($\mathfrak{H}$) Paul Alexandroff und Paul Urysohn, „Zur Theorie der topologischen Räume".
($\mathfrak{B}$) Paul Alexandroff, „Über die Struktur der bikompakten topologischen Räume".
($\mathfrak{S}$) Paul Urysohn, „Über die Metrisation der kompakten topologischen Räume".
Diese Arbeiten werden kurz durch $\mathfrak{H}$, $\mathfrak{B}$, $\mathfrak{S}$ zitiert.

Der Begriff des im Kleinen bikompakten Raumes scheint mir von sehr großer topologischer Wichtigkeit zu sein — man denke etwa an den gewöhnlichen Euklidischen Raum, an ein ein-, zwei- oder mehrdimensionales Gebiet, an eine (insbesondere abstrakt gefaßte) offene Riemannsche Fläche[2]): alle diese seit Zeiten klassisch gewordenen fundamentalen topo-

[2]) Die erste abstrakte Fassung des Flächenbegriffes scheint von Weyl („*Die Idee der Riemannschen Fläche*") zu stammen. Bis auf ein unwesentliches Versehen*) stimmt der Grundgedanke der Weylschen Flächendefinition mit der folgenden Erklärung überein:

Ein topologischer Raum heißt eine Fläche, falls eine gewisse Umgebung jedes seiner Punkte ein umkehrbar eindeutiges und umkehrbar stetiges [im Sinne des Raumes**)] *Bild des Inneren eines gewöhnlichen Kreises ist.*

Weyl nennt weiter eine Fläche *geschlossen*, falls sie, als topologischer Raum betrachtet, kompakt ist. Nun sieht man sofort, daß diese Definition sich auf von dem anschaulichen Sinne des Begriffes der geschlossenen Fläche sehr entfernte Bildungen anwenden läßt. Man betrachtet, um ein Beispiel dafür zu erhalten, z. B. eine geordnete Menge Θ, die den Ordnungstypus

$$(\lambda+1)\,\Omega^* + \lambda + (1+\lambda)\,\Omega$$

hat***).

Man denkt sich ferner die Menge F, die aus allen möglichen Paaren (x, y) von Elementen von Θ gebildet ist. Diese Menge F wird zu einer im Weylschen Sinne geschlossenen „Fläche" $\mathfrak{F}$, wenn man die Umgebungen $U(\xi)$ ihrer sämtlichen „Punkte" $\xi = (x, y)$ folgendermaßen definiert:

$$U(\xi) = U((x, y)) = \text{Inbegriff aller } \tilde{\xi}(\tilde{x}, \tilde{y}), \quad \text{wo} \quad \begin{cases} x_1 < \tilde{x} < x_2 \\ y_1 < \tilde{y} < y_2 \end{cases} \quad (\text{in } \Theta)$$

und $x_1, y_1;\ x_2, y_2$ zwei beliebige den Ungleichungen $x_1 < x < x_2$, $y_1 < y < y_2$ genügende Elementenpaare der Menge Θ sind. Die Fläche $\mathfrak{F}$, die die Gestalt einer nicht-Archimedischen Ebene hat, entspricht aber kaum dem anschaulichen Wesen einer geschlossenen Fläche.

Dieser Übelstand wird vermieden, indem man statt der Kompaktheit die Bikompaktheit als charakteristische Eigenschaft der geschlossenen Flächen verlangt. Dieses Verfahren scheint mir noch den folgenden Vorzug zu haben. Zufolge dem Satze III (§ 3 dieser Arbeit) ergibt es sich sofort, daß eine in meinem Sinne geschlossene Fläche stets dem Hausdorffschen II. Abzählbarkeitsaxiom genügt, und folglich nach einem sehr bemerkenswerten Satze von Herrn Urysohn als ein metrischer Raum (ebenfalls im Hausdorffschen Sinne) erklärt sein kann. Außerdem läßt sich, zwar durch recht umständliche Überlegungen, aber ohne irgendwelche prinzipielle Schwierigkeit, die Triangulationsfähigkeit einer jeden (in meinem Sinne) geschlossenen Fläche auf Grund rein topologischer Begriffsbildungen beweisen.

*) Das, soweit ich es beurteilen kann, darin liegt, daß aus den Weylschen Umgebungsaxiomen im allgemeinen keine Möglichkeit folgt, zwei verschiedene Punkte einer Fläche durch zueinander fremde Umgebungen (im Sinne des Hausdorffschen Axiomes (D), loc. cit. S. 213) zu trennen.

**) Über stetige Abbildungen der topologischen Räume vgl. Hausdorff, loc. cit. S. 358 ff.

***) Ω ist die erste unabzählbare Ordnungszahl, Ω^* der dazu inverse Ordnungstypus, λ der Ordnungstypus der reellen Zahlen. Vgl. Hausdorff, S. 73, 93, 125 ($\Omega = \omega_1$).

logischen Bildungen sind, als Räume betrachtet, nicht kompakt, wohl aber im Kleinen bikompakt. Anderseits sind mir nur wenige topologische Sätze bekannt, die für ein weiteres Gebiet als die im Kleinen bikompakten Räume gültig bleiben. Die soeben erwähnten Räume scheinen mir deshalb einen aus sachlichen Gründen naturgemäßen Spielraum der allgemein-topologischen Erscheinungen zu bilden.

2. Allgemeine Eigenschaften der im Kleinen kompakten Räume.

Satz I. *In jedem Punkte ξ, in welchem ein gegebener topologischer Raum bikompakt ist, stimmt sein Durchschnittscharakter $\psi_{\mathfrak{R}}(\xi)$ mit seinem Charakter $\chi_{\mathfrak{R}}(\xi)$ überein; außerdem ist unter diesen Voraussetzungen der Konvergenzcharakter $\varkappa_{\mathfrak{R}}(\xi)$ stets definiert, und zwar ist $\varkappa_{\mathfrak{R}}(\xi) \leqq \chi_{\mathfrak{R}}(\xi)$; endlich ist der Punkt ξ ein regulärer Punkt des Raumes*[3]).

Dieser Satz folgt unmittelbar aus den für die bikompakten Räume schon bewiesenen Eigenschaften.

Fundamentalsatz 1. *Ein jeder im Kleinen bikompakte topologische Raum $\mathfrak{R}$ läßt sich* (falls er nicht selbst bikompakt ist) *durch Hinzufügung eines einzigen Punktes zu einem bikompakten Raume vervollständigen; dies ist außerdem nur auf eine Weise möglich.*

Wir nennen „β-Gebiet" ein solches Gebiet Γ, daß die abgeschlossene Menge $\overline{\Gamma}$, als Relativraum betrachtet, bikompakt ist.

Es ist sofort klar, daß der gegebene Raum $\mathfrak{R}$, falls er nicht bikompakt ist, in keiner endlichen Menge von abgeschlossenen Mengen $\overline{\Gamma}$ enthalten ist, dagegen ist jeder Punkt des Raumes $\mathfrak{R}$ (da $\mathfrak{R}$ im Kleinen bikompakt ist) immer wenigstens in einem β-Gebiete enthalten. Das Gebiet $\mathfrak{G} = \mathfrak{R} - (\overline{\Gamma}_1 + \overline{\Gamma}_2 + \ldots + \overline{\Gamma}_n)$ ist also immer von Null verschieden, wie auch die β-Gebiete $\Gamma_1, \Gamma_2, \ldots, \Gamma_n$ gewählt seien. Nun fügen wir dem Raume $\mathfrak{R}$ einen neuen Punkt ξ hinzu mittels der Umgebungen $U(\xi) = \xi + \mathfrak{G}$. Man überzeugt sich sofort, daß in $\mathfrak{R} + \xi$ die Hausdorffschen Umgebungsaxiome erfüllt sind; außerdem strömt[4]) jede unendliche, im Raume $\mathfrak{R}$ keinen vollständigen Häufungspunkt besitzende Menge E im Raume $\mathfrak{R} + \xi$ dem Punkte ξ zu. $\mathfrak{R} + \xi$ ist infolgedessen ein bikompakter topologischer Raum. Es gelingt auch zu zeigen, ohne besondere Schwierigkeit, daß ein jedes andere Umgebungssystem für einen adjungierten Punkt ξ dem soeben konstruierten gleichwertig ist, solange nur der Raum $\mathfrak{R} + \xi$ bikompakt bleiben soll, und damit wird der Fundamentalsatz 1 vollständig bewiesen.

Anmerkungen. Die soeben ausgesprochene Eindeutigkeitseigenschaft ist darum besonders bemerkenswert, weil man im allgemeinen einen im

[3]) Vgl. für die Bezeichnungen die schon erwähnten Arbeiten $\mathfrak{H}$ und $\mathfrak{G}$. Daselbst sind die entsprechenden Sätze für bikompakte Räume schon bewiesen worden.

Kleinen bikompakten Raum durch Hinzufügung eines einzigen Punktes auf unendlich viele Arten zu einem absolut abgeschlossenen (irregulären[4])) Raume erweitern kann.

Wegen des Satzes über die Mächtigkeit der perfekten Mengen in bikompakten Räumen[4]) ergibt sich der analoge Satz für im Kleinen bikompakte Räume als eine unmittelbare Folge des soeben ausgesprochenen Fundamentalsatzes.

Der Fundamentalsatz läßt sich umkehren; es besteht offenbar auch der Satz:

Jedes in einem bikompakten Raume liegende Gebiet ist im Kleinen bikompakt (als Relativraum betrachtet).

Die Menge aller im Kleinen bikompakten Räume ist daher mit der Menge der in bikompakten Räumen liegenden Gebiete identisch[5]).

Wenn man das Wort „bikompakt" durch „kompakt" ersetzt, so erhält man einen dem Fundamentalsatz 1 *vollständig analogen Satz.* Der umgekehrte Satz ist auch richtig, d. h. jedes aus einem kompakten Raume durch Weglassung eines einzigen Punktes entstehende Gebiet ist, als Raum betrachtet, im Kleinen kompakt.

Es gibt aber in kompakten Räumen liegende Gebiete, die nicht im Kleinen kompakt sind.

Es ist noch zu bemerken, daß man durch Adjunktion eines einzigen Punktes zu einem im Kleinen kompakten Raume allgemein mehrere verschiedene Räume erhalten kann (für die Eindeutigkeitseigenschaft ist daher die Voraussetzung der Bikompaktheit unentbehrlich).

3. Das II. Abzählbarkeitsaxiom und das Metrisationsproblem der topologischen Räume.

Das allgemeine Metrisationsproblem der topologischen Räume besteht darin, die Bedingungen zu finden, die notwendig und hinreichend sind, damit ein topologischer Raum einem metrischen topologisch identisch (= homöomorph) ist. Herr Urysohn hat bereits das Problem der Metrisation der kompakten Räume gelöst; er bewies nämlich, daß in diesem

[4]) Vgl. die vorige Fußnote.

[5]) Es ist zu bemerken, daß außerdem jede in einem bikompakten Raume $\mathfrak{R}$ liegende abgeschlossene Menge F durch einen einzigen Punkt ξ (mit passend gewählten Umgebungen) so ersetzt werden kann, daß damit ein wiederum bikompakter Raum $\mathfrak{R}_1$ entsteht, dessen Gebiet $\mathfrak{R}_1 - \xi$ mit dem ursprünglichen Gebiete $\mathfrak{R} - F$ identisch ist.

Falle die gewünschte Bedingung im II. Abzählbarkeitsaxiome besteht[6]). Da dasselbe Axiom eine fundamentale Bedeutung auch für das im vorliegenden Aufsatze von mir gelöste Metrisationsproblem der im Kleinen kompakten Räume hat, so halte ich es für zweckmäßig, auf den Zusammenhang zwischen diesem Axiome und anderen naheliegenden Raumeigenschaften kurz hinzuweisen:

Satz II. *In einem metrischen Raume ist eine jede der folgenden Eigenschaften*[7]) *dem II. Abzählbarkeitsaxiome äquivalent:*

a) *Es gibt keine unabzählbare Menge paarweise fremder Gebiete;*

b) *jede wohlgeordnete ab- oder zunehmende Menge von verschiedenen abgeschlossenen Mengen (oder Gebieten) ist höchstens abzählbar;*

c) *jede abgeschlossene Menge zerfällt in einen perfekten und in einen höchstens abzählbaren Bestandteil;*

d) *die Summe von beliebig vielen Gebieten kann stets durch eine Summe von höchstens abzählbar vielen Gebieten ersetzt werden;*

e) *es gibt eine abzählbare überall dichte Teilmenge.*

Dagegen gibt es (sogar bikompakte) *topologische* Räume, in denen alle diese Eigenschaften sowie das erste Abzählbarkeitsaxiom gleichzeitig erfüllt sind, das II. Abzählbarkeitsaxiom aber ungültig ist.

Satz III. *Wenn in einem bikompakten topologischen Raume $\mathfrak{R}$ das I., aber nicht das II. Abzählbarkeitsaxiom gültig ist, dann gibt es eine perfekte Teilmenge $P \subset \mathfrak{R}$ von der Eigenschaft, daß keine Umgebung irgendeines Punktes von P, als Relativraum betrachtet, demselben Axiome genügt*[8]).

Die perfekte Menge P kann sowohl nirgends dicht im Raume liegen, als auch den ganzen Raum (oder ein gewisses Teilgebiet) vollständig erfüllen.

Satz IV. *Damit ein regulärer, dem II. Abzählbarkeitsaxiome genügender Raum $\mathfrak{R}$ kompakt sei, ist es notwendig und hinreichend, daß ein jeder diese beiden Eigenschaften besitzende und den Raum $\mathfrak{R}$ als überall dichte Teilmenge enthaltende topologische Raum R mit $\mathfrak{R}$ identisch ist.*

---

[6]) Diese Annalen **92**, S. 275–293. (Diese Arbeit ist bereits durch „$\mathfrak{S}$" zitiert.) — Während der Drucklegung dieser Arbeit ist eine gemeinschaftliche Note von P. Urysohn und mir in den Pariser Comptes Rendus **177**, S. 1274 erschienen, wo das allgemeine Metrisationsproblem für die topologischen Räume gelöst ist. Übrigens ist zu bemerken, daß die Einbettung des hier behandelten Falles in das allgemeine Metrisationskriterium nicht einfacher ist als der hier wiedergegebene Beweis selbst. Dasselbe gilt auch vom soeben erwähnten Urysohnschen Satze.

[7]) Daß alle diese Eigenschaften in topologischen Räumen aus dem II. Abzählbarkeitsaxiome folgen, war schon von Hausdorff bewiesen (loc. cit., Kap. VIII, § 3).

[8]) Für die kompakten (nicht bikompakten) Räume ist der Satz im allgemeinen falsch.

Der Beweis der Sätze II und III ist leicht; wegen des Beweises des Satzes IV verweise ich auf die in den „*Fundamenta Mathematicae*" erscheinenden Arbeiten[9]), wo die ganze Theorie ausführlich dargestellt ist.

4. Lösung des Metrisationsproblemes für die im Kleinen kompakten topologischen Räume.

Aus dem Fundamentalsatze 1 und dem Satze $\mathfrak{G}$, I ergibt sich sofort, daß *jeder im Kleinen kompakte, dem II. Abzählbarkeitsaxiome genügende Raum* $\mathfrak{R}$ *metrisierbar ist.* Tatsächlich läßt sich ja $\mathfrak{R}$ zu einem bikompakten Raume $\mathfrak{R} + \xi = R$ erweitern; zufolge dem in $\mathfrak{R}$ erfüllten II. Abzählbarkeitsaxiome und der Regularität des Raumes ist der Punkt ξ der einzige Durchschnittspunkt einer gewissen abzählbaren Menge von Gebieten, also ist nach dem Satze $\mathfrak{G}$, I der Charakter des Raumes R im Punkte ξ gleich $\aleph_0$, also bekommt man ein abzählbares Umgebungssystem für den gesamten Raum R, indem man zu den (abzählbar vielen) Umgebungen der sämtlichen Punkte von $\mathfrak{R}$ noch die Umgebungen des Punktes ξ nimmt. Da R bikompakt ist, so ist R zufolge dem Urysohnschen Satze auch metrisierbar; $\mathfrak{R}$ ist also nicht nur metrisierbar, sondern der ganze Raum $\mathfrak{R}$ liegt als Gebiet in einem kompakten metrischen Raume. Das II. Abzählbarkeitsaxiom drückt also eine hinreichende, aber (wie man sich durch leichte Gegenbeispiele überzeugt) keine notwendige Bedingung für die Metrisationsfähigkeit eines im Kleinen kompakten topologischen Raumes aus. Es besteht dagegen der

Fundamentalsatz 2. *Damit ein im Kleinen kompakter topologischer Raum metrisierbar sei, ist es notwendig und hinreichend, daß der Raum entweder dem II. Abzählbarkeitsaxiom genügt oder sich in eine Menge* (beliebiger Mächtigkeit) *von paarweise fremden Gebieten zerspalten läßt, von denen ein jedes* (als Relativraum betrachtet) *dem II. Abzählbarkeitsaxiom genügt*[10]).

Es bedarf kaum eines Beweises, daß die soeben ausgesprochene Bedingung eine hinreichende ist. Man braucht also nur zu zeigen, daß jeder im Kleinen kompakte metrische[11]) Raum, falls er dem II. Abzählbarkeitsaxiom nicht genügt, in eine Menge von diesem Axiom genügenden Gebieten zerfällt; das geschieht aber wie folgt:

---

[9]) P. Alexandroff u. P. Urysohn, Mémoire sur les espaces topologiques compacts; P. Alexandroff, Sur les espaces localement compacts.

[10]) Der Raum ist in diesem Fall auch bikompakt im Kleinen.

[11]) Ich schreibe schlechthin „*metrischer Raum*" statt „*metrisierbarer topologischer Raum*"; da in dieser Arbeit nur topologische Eigenschaften von metrischen Räumen vorkommen, so scheint mir kein Mißverständnis möglich zu sein.

Hilfssatz. *Vorausgesetzt sei ein metrischer Raum, der im Kleinen dem II. Abzählbarkeitsaxiom genügt*[12]); *dann läßt sich ein System von den ganzen Raum vollständig definierenden, sphärischen Umgebungen derart auswählen, daß jeder Raumpunkt in höchstens abzählbar vielen dieser Umgebungen enthalten ist.*

Mit Zuhilfenahme des Zermeloschen Wohlordnungssatzes denken wir uns die Menge aller derjenigen sphärischen Umgebungen des gegebenen Raumes $\mathfrak{R}$, in welchen irgendeine abzählbare Menge überall dicht ist, in eine wohlgeordnete Menge

$$V_1 V_2 \ldots V_n \ldots V_\alpha \ldots$$

von einem gewissen Ordnungstypus Ω_τ (wo Ω_τ eine Anfangszahl ist) verwandelt. Wegen der Voraussetzung gibt es in V_1 eine überall dichte, höchstens abzählbare Menge D_1. Wir setzen noch identisch $G_1 = V_1$. Nehmen wir an, es seien G_α und D_α für alle $\alpha < \lambda$ konstruiert, so setzen wir definitionsgemäß

$$G_\lambda = V_\lambda - \overline{(\sum_{\alpha<\lambda} G_\alpha)}$$

und es sei dann D_λ eine bestimmte, höchstens abzählbare, in G_λ überall dichte Menge. Wir definieren endlich

$$D = \sum_{\alpha<\Omega_\tau} D_\alpha.$$

Es läßt sich leicht beweisen, daß

1. D in $\mathfrak{R}$ überall dicht ist;

2. die Mächtigkeit von der Durchschnittsmenge $D \cdot V_\alpha$ für jedes α kleiner oder gleich $\aleph_0$ ist.

Es sei nun x ein beliebiger Raumpunkt und ϱ_x die obere Grenze aller positiven Zahlen r_x von der Art, daß ein jedes $S(x, r_x)$ [13]) eine höchstens abzählbare, überall dichte Teilmenge enthält. Das System Σ von allen Sphären $S(x, r)$, wo $x \subset D$ und $r < \frac{1}{8}\varrho_x$ eine positive rationale Zahl ist, bildet dann, wie man ohne besondere Schwierigkeiten beweisen kann, ein Umgebungssystem, wie wir es haben wollten[14]). Damit wird der Hilfssatz bewiesen.

Es sei nun ξ ein beliebiger Punkt unseres Raumes, $\Gamma_0^{(\xi)}$ irgendeine Sphäre des Systems Σ, die den Punkt ξ enthält.

---

[12]) Das heißt eine gewisse Umgebung eines jeden Punktes soll, als topologischer Raum betrachtet, dem II. Abzählbarkeitsaxiom genügen.

[13]) $S(x, r_x)$ = Sphäre vom Radius r_x und mit dem Zentrum x (im beliebigen metrischen Raume ist diese Sphäre der Inbegriff aller von x weniger als um r_x entfernten Punkte).

[14]) Der wesentliche Punkt des Beweises liegt in der Tatsache, daß aus $\xi \subset S(x, r)$ wo $S(x, r)$ eine Umgebung des Systems Σ ist, die Relation $x \subset S(\xi, \varrho_x)$ folgt.

Es sei $\Gamma_n^{(\xi)}$ schon definiert, und zwar so, daß dabei das II. Abzählbarkeitsaxiom gültig ist (bei $\Gamma_0^{(\xi)}$ ist es ja der Fall). Wir definieren $\Gamma_{n+1}^{(\xi)}$ als Summe von $\Gamma_n^{(\xi)}$ und allen Sphären des Systems Σ, deren Durchschnitt mit $\Gamma_n^{(\xi)}$ von Null verschieden ist. Man sieht sofort[15]), daß $\Gamma_{n+1}^{(\xi)}$ und also auch

$$\Gamma^{(\xi)} = \Gamma_0^{(\xi)} + \Gamma_1^{(\xi)} + \Gamma_2^{(\xi)} + \ldots + \Gamma_n^{(\xi)} + \ldots$$

dem II. Abzählbarkeitsaxiom genügt. Jetzt sagen wir für einen Augenblick, daß der Punkt x mit dem Punkte ξ verbunden ist, falls eine endliche Anzahl von Sphären $S_1, S_2, \ldots, S_\varkappa$ des Systems Σ derart existiert, daß

$$1.\quad S_1 \supset \xi; \qquad 2.\quad S_i \cdot S_{i+1} \neq 0; \qquad 3.\quad S_\varkappa \supset x \qquad (1 \leqq i \leqq \varkappa - 1).$$

Dann beweist man sofort, daß $\Gamma^{(\xi)}$ aus allen mit ξ verbundenen Punkten des Raumes $\mathfrak{R}$ gebildet ist. Infolgedessen ist es leicht zu sehen, daß, wenn ξ und η verschiedene Raumpunkte sind, entweder $\Gamma^{(\xi)} = \Gamma^{(\eta)}$ oder $\Gamma^{(\xi)} \cdot \Gamma^{(\eta)} = 0$ ist. Die Zerfällung des Raumes in punktfremde, dem II. Abzählbarkeitsaxiom genügende Gebiete ist damit erbracht, und der Fundamentalsatz 2 bewiesen.

Korollar 1. *Damit ein zusammenhängender*[16]), *im Kleinen kompakter topologischer Raum metrisierbar sei, ist das II. Abzählbarkeitsaxiom gleichzeitig notwendig und hinreichend.*

Korollar 2. *Falls in einem zusammenhängenden metrischen Raume das II. Abzählbarkeitsaxiom nicht gültig ist, existiert wenigstens ein Punkt, so daß in keiner Umgebung desselben jenes Axiom erfüllt ist.*

Alle diese Sätze können durch leicht konstruierbare Beispiele veranschaulicht werden; sie erweisen sich außerdem als keiner weiteren Verschärfung oder Verallgemeinerung fähig.

Göttingen, den 10. Juli 1923.

P. S. Die Hauptergebnisse dieser Arbeit sind im Jahre 1922 der Moskauer Mathematischen Gesellschaft und am 26. Juni 1923 der Göttinger Mathematischen Gesellschaft mitgeteilt worden.

---

[15]) Das ist eben die wesentliche Folge der durch den Hilfssatz ausgedrückten Eigenschaft des Umgebungssystems Σ.

[16]) Im Hausdorffschen Sinne, loc. cit. S. 244.

(Eingegangen am 1. 8. 1923.)

Der Hilbertsche Raum als Urbild der metrischen Räume*).

Von

Paul Urysohn † in Moskau.

Der Zweck dieser Arbeit ist, zu zeigen, daß die Begriffe „separable (D)-Menge" (im Fréchetschen Sinne) und „Teilmenge des Hilbertschen Raumes[1])" topologisch identisch sind; oder, anders ausgedrückt, daß *die notwendige und hinreichende Bedingung, damit ein metrischer Raum*[2]) *E einer Teilmenge des Hilbertschen Raumes H homöomorph sei, darin besteht, daß E eine abzählbare dichte Teilmenge besitzt.*

Die Notwendigkeit dieser Bedingung ist unmittelbar einleuchtend[3]); es handelt sich also um den Beweis folgenden Satzes:

Jeder metrische Raum mit abzählbarer dichter Teilmenge (*also insbesondere jeder kompakte metrische Raum*[4])) *ist einer Teilmenge des Hilbertschen Raumes homöomorph.*

Vorbemerkung. Der Hilbertsche Raum ist *selbst* nicht kompakt, besitzt aber in sich kompakte *Quadern*, z. B. den durch die Ungleichungen

$$0 \leqq x_n \leqq \frac{1}{n} \quad {}^{5})$$

definierten Quader Q. Q kann also als ein kompakter metrischer Raum betrachtet werden. Wir werden nunmehr beweisen, daß jeder unseren Voraussetzungen genügender Raum E *einer Teilmenge des Quaders Q homöomorph ist*; daraus folgt insbesondere, daß jeder metrische Raum mit

*) Vgl. meine inzwischen erschienene Note „Les classes (D) séparables ...", Comptes Rendus Paris **178**, S. 65.

[1]) Die Definition des Hilbertschen Raumes kann man z. B. in Hausdorffs *Grundzügen der Mengenlehre*, S. 287, IV finden.

[2]) Hausdorff, l. c. S. 211.

[3]) Man vergleiche hierzu Hausdorff, l. c. S. 287, 288 und 273, VIII.

[4]) l. c. S. 274, X.

[5]) $x = (x_1, x_2, \ldots, x_n, \ldots)$ ist ein Punkt von H.

abzählbarer dichter Teilmenge vom topologischen Standpunkte aus als Teil eines *kompakten* metrischen Raumes betrachtet werden kann.

Beweis. Bekanntlich ist jeder metrische Raum einem *beschränkten* Raume homöomorph[6]); es genügt also nur den Fall zu betrachten, wo die *Breite*[7]) von $E \leqq 1$ ist. Es sei ferner

$$E_0 = \{a_1, a_2, \ldots, a_n, \ldots\}$$

eine in E dichte abzählbare Menge, die wir uns fest gewählt denken. Wenn ξ ein beliebiger Punkt von E ist, und wenn wir die Entfernungen

$$\varrho(\xi, a_1), \quad \varrho(\xi, a_2), \quad \ldots, \quad \varrho(\xi, a_n), \ \ldots$$

der Reihe nach mit $x_1, 2x_2, \ldots, n x_n, \ldots$, bezeichnen, so ordnen wir dem Punkte ξ von E den Punkt $x = (x_1, x_2, \ldots, x_n, \ldots)$ von H zu. Da $x_n = \frac{1}{n}\varrho(\xi, a_n) \leqq \frac{1}{n}$ ist, so wird damit E auf eine Teilmenge M des Quaders Q abgebildet.

Diese Abbildung ist eineindeutig. Wenn in der Tat ξ und η zwei verschiedene Punkte von E sind, so können wir wegen der Dichtigkeit von E_0 in E einen der Bedingung

$$\varrho(\xi, a_n) < \frac{1}{2}\varrho(\xi, \eta)$$

genügenden Punkt a_n finden; folglich ist

$$\varrho(\xi, a_n) < \frac{1}{2}\varrho(\xi, a_n) + \frac{1}{2}\varrho(a_n, \eta),$$

also

$$\varrho(\xi, a_n) < \varrho(\eta, a_n),$$

d. h. $x_n < y_n$ [8]) und $x \neq y$.

Wir zeigen jetzt, daß *die erhaltene Abbildung nebst ihrer Umkehrung stetig ist.*

1. Wenn wir zur Abkürzung $\varrho(\xi, \eta)$ mit λ bezeichnen, so ist für jedes n

$$|\varrho(\xi, a_n) - \varrho(\eta, a_n)| \leqq \varrho(\xi, \eta) = \lambda,$$

also

$$|x_n - y_n| \leqq \frac{\lambda}{n},$$

$$\varrho(x, y) = \sqrt{\sum_{n=1}^{\infty}(x_n - y_n)^2} \leqq \sqrt{\lambda^2 \sum_{n=1}^{\infty}\frac{1}{n^2}} = \frac{\lambda\pi}{\sqrt{6}} < 2\lambda,$$

$$\varrho(x, y) < 2\varrho(\xi, \eta),$$

woraus sogar die gleichmäßige Stetigkeit der Abbildung von E auf M folgt.

---

[6]) l. c. S. 312.

[7]) l. c. S. 290.

[8]) y_n ist die n-te Koordinate des dem Punkte η zugeordneten Punktes y.

2. Es sei jetzt x irgendein Punkt von M, ξ der entsprechende Punkt von E. Es bleibt zu zeigen, daß es zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart gibt, daß $\varrho(x, y) < \delta$ die Ungleichung $\varrho(\xi, \eta) < \varepsilon$ zur Folge hat.

Zu jedem ε können wir aber einen der Ungleichung $\varrho(\xi, a_n) < \frac{\varepsilon}{3}$ genügenden Punkt a_n wählen[9]); es genügt alsdann $\delta = \frac{\varepsilon}{3n}$ zu setzen: es folgt dann in der Tat aus $\varrho(x, y) < \delta$, daß

$$\frac{\varepsilon}{3n} = \delta > \varrho(x, y) = \sqrt{\sum_{n=1}^{\infty} (x_n - y_n)^2} \geqq y_n - x_n = \frac{1}{n}\varrho(\eta, a_n) - \frac{1}{n}\varrho(\xi, a_n),$$

also

$$\varrho(\xi, \eta) \leqq \varrho(\xi, a_n) + \varrho(a_n, \eta) < \varrho(\xi, a_n) + \varrho(\xi, a_n) + \frac{\varepsilon}{3} < \varepsilon$$

ist, w. z. b. w.

Daher ist die Homöomorphie von E und M bewiesen.

Göttingen, 22. Juli 1923.

---

[9]) n hängt von x und ε ab.

(Eingegangen am 1. 8. 1923.)

Über die Bausteine der mathematischen Logik.

Von

M. Schönfinkel in Moskau[1]).

§ 1.

Es entspricht dem Wesen der axiomatischen Methode, wie sie heute vor allem durch die Arbeiten Hilberts zur Anerkennung gelangt ist, daß man nicht allein hinsichtlich der Zahl und des Gehalts der *Axiome* nach möglichster Beschränkung strebt, sondern auch die Anzahl der als undefiniert zugrunde zu legenden *Begriffe* so klein wie möglich zu machen sucht, indem man nach Begriffen fahndet, die vorzugsweise geeignet sind, um aus ihnen alle anderen Begriffe des fraglichen Wissenszweiges aufzubauen. Begreiflicherweise wird man sich im Sinne dieser Aufgabe bezüglich des Verlangens nach Einfachheit der an den Anfang zu stellenden Begriffe entsprechend bescheiden müssen.

Bekanntlich lassen sich die grundlegenden *Aussagenverknüpfungen* der mathematischen Logik, die ich hier in der von Hilbert in seinen Vorlesungen verwendeten Bezeichnungsweise wiedergebe:

$$\bar{a}, \quad a \vee b, \quad a \,\&\, b, \quad a \rightarrow b, \quad a \sim b$$

(lies: „a nicht", „a oder b", „a und b", „wenn a, so b", „a äquivalent b"), aus einer einzigen von ihnen überhaupt nicht, aus zweien von ihnen aber nur in der Weise gewinnen, daß man die Negation und irgendeine der drei folgenden Verknüpfungen als undefiniert zugrunde legt. (Von diesen drei Arten der Zurückführung haben Whitehead und Russell die erste und Frege die dritte verwendet.)

Daß dessenungeachtet die Zurückführung auf eine einzige Grundverknüpfung sehr wohl möglich ist, sobald man sich von der Einschränkung

[1]) Die folgenden Gedanken wurden vom Verfasser am 7. Dez. 1920 vor der Mathematischen Gesellschaft in Göttingen vorgetragen. Ihre formale und stilistische Durcharbeitung für diese Veröffentlichung wurde von H. Behmann in Göttingen übernommen.

frei macht, daß diese gerade der obigen Reihe entnommen sein soll, ist erst neuerdings von Sheffer[2]) entdeckt worden. Wählt man nämlich als Grundverknüpfung etwa „a nicht oder b nicht", d. h. „von den Aussagen a und b ist mindestens eine falsch", die mit den obigen Zeichen in den beiden äquivalenten Formen

$$\bar{a} \vee \bar{b} \quad \text{und} \quad \overline{a \,\&\, b}$$

geschrieben werden kann, und als zugehöriges neues Zeichen

$$a \mid b,$$

so ist augenscheinlich

$$\bar{a} = a \mid a, \qquad a \vee b = (a \mid a) \mid (b \mid b),$$

womit wegen

$$a \,\&\, b = \overline{\bar{a} \vee \bar{b}}, \qquad (a \rightarrow b) = \bar{a} \vee b, \qquad (a \sim b) = (a \rightarrow b) \,\&\, (b \rightarrow a)$$

die Zurückführung grundsätzlich geleistet ist.

Es ist nun bemerkenswerterweise sogar noch darüber hinaus möglich, durch eine geeignete Abänderung der Grundverknüpfung auch die beiden höheren Aussagen

$$(x) f(x) \quad \text{und} \quad (Ex) f(x),$$

d. h. „Alle Individuen haben die Eigenschaft f" und „Es gibt ein Individuum, das die Eigenschaft f hat", mit andern Worten, die beiden Operationen (x) und (Ex), die mit den früheren zusammen bekanntlich ein im Sinne der Axiomatik vollständiges System von Grundverknüpfungen der mathematischen Logik ausmachen, mit zu erfassen.

Verwenden wir nämlich als Grundbeziehung nunmehr

$$(x)[\overline{f(x)} \vee \overline{g(x)}] \quad \text{bzw.} \quad (x)\overline{f(x) \,\&\, g(x)}$$

und schreiben wir hierfür

$$f(x) \mid^{x} g(x),$$

so gilt offenbar (da wir Konstante formal wie Funktionen eines Arguments behandeln dürfen):

$$\bar{a} = a \mid^{x} a, \qquad a \vee b = (x)(\bar{\bar{a}} \vee \bar{\bar{b}}) = \bar{a} \mid^{x} \bar{b} = (a \mid^{y} a) \mid^{x} (b \mid^{y} b),$$

$$(x) f(x) = (x)(\overline{\overline{f(x)}} \vee \overline{\overline{f(x)}}) = \overline{f(x)} \mid^{x} \overline{f(x)} = (f(x) \mid^{y} f(x)) \mid^{x} (f(x) \mid^{y} f(x)),$$

womit wegen

$$(Ex) f(x) = \overline{(x)\overline{f(x)}}$$

auch die neue Behauptung erwiesen ist.

Die bisherigen Erfolge auf dem eingeschlagenen Wege ermutigen zu dem Versuch eines weiteren Fortschreitens. Man wird auf den im ersten

[2]) Am. Math. Soc. Trans. **14** (1913), S. 481.

Augenblick gewiß äußerst kühn erscheinenden Gedanken geführt, hinsichtlich völlig beliebiger logisch allgemeiner Aussagen — für andere hätte das Verlangen offenbar keinen Sinn — auch die noch verbleibenden Grundbegriffe der Aussage, der Aussagenfunktion und der Veränderlichen durch geeignete Zurückführung zu beseitigen zu suchen. Eine derartige Möglichkeit näher zu prüfen und zu verfolgen, wäre nicht nur vom methodischen Standpunkt des Strebens nach größtmöglicher gedanklicher Einheitlichkeit, sondern auch von einem gewissen philosophischen oder, wenn man will, ästhetischen Standpunkt wertvoll, insofern nämlich, als die Veränderliche in der logischen Aussage ja nichts weiter als ein Abzeichen ist, um gewisse Argumentstellen und Operatoren als zusammengehörig zu kennzeichnen, und somit den Charakter eines bloßen, dem konstanten, „ewigen" Wesen der logischen Aussage eigentlich unangemessenen Hilfsbegriffes hat.

Es erscheint mir äußerst bemerkenswert, daß auch das eben aufgestellte Ziel der Verwirklichung fähig ist, und zwar in dem Sinne, daß hier die Zurückführung auf drei Grundzeichen gelingt.

§ 2.

Zur Erreichung dieser letzten und tiefsten Zurückführung bedarf es nun freilich einer Reihe von Hilfsmitteln und Tatsachenzusammenhängen, die zunächst bereitgestellt und erklärt werden müssen.

Es wird darum notwendig sein, unser Problem an dem vorhin erreichten Punkte stehen zu lassen und vorerst eine Art von *Funktionenkalkül* — in einem gegenüber dem sonst üblichen verallgemeinerten Sinne — zu entwickeln.

Unter einer *Funktion* versteht man bekanntlich im einfachsten Falle eine Zuordnung zwischen den Elementen irgendeines Bereiches von Größen, des Argumentbereiches, und denen eines (zumeist freilich mit dem ersten zusammenfallend gedachten) Bereiches der Funktionswerte, derart, daß jedem Argumentwert höchstens ein Funktionswert entspricht. Dieser Begriff soll hier nun in dem Sinne erweitert werden, daß sowohl als Argumentwerte wie auch als Funktionswerte selbst wieder Funktionen auftreten können. Den Wert einer Funktion f für den Argumentwert x bezeichnen wir durch einfaches Nebeneinanderstellen des Funktions- und des Argumentzeichens, d. h. als

$$fx.$$

Funktionen mehrerer Argumente lassen sich auf Grund unserer erweiterten Begriffsbestimmung der Funktion in der folgenden Weise auf solche eines Arguments zurückführen:

Wir fassen z. B.

$$F(x, y)$$

etwa als eine Funktion des einzigen Arguments y, aber nicht mehr als eine fest gegebene, sondern als eine veränderliche, ihrer Gestalt nach von x abhängige Funktion auf. (Es handelt sich hier — wohlverstanden — um eine Abhängigkeit der *Funktion*, also der Zuordnung selbst, nicht etwa um die selbstverständliche Abhängigkeit des Funktions*wertes* vom Argument.) In der Mathematik pflegt man hier von einer Funktion zu reden, die noch von einem Parameter abhängt, und etwa

$$G_x(y)$$

zu schreiben. Wir können diese Funktion G selbst — ihre Gestalt sozusagen — als den Wert (Funktionswert) einer neuen Funktion f betrachten, so daß $G = fx$ ist.

Wir schreiben daher in unserer Symbolik

$$(fx)y$$

oder, indem wir, wie dies z. B. aus der Theorie der unendlichen Reihen geläufig ist, verabreden, daß Klammern, die das linke Ende einer derartigen symbolischen Form mit umfassen, wegbleiben dürfen, einfacher

$$fxy,$$

wo die neue Funktion f von der früheren F wohl zu unterscheiden ist.

Ich möchte die eben geschilderte Umformung dadurch dem Verständnis näherbringen, daß ich sie auf die spezielle Zahlenfunktion $x - y$ anwende. Betrachtet man den Ausdruck als Funktion von y allein, so hat diese die „Gestalt" $x -$, bedeutet also die „Differenz des x mit irgendeiner gegebenen Größe", womit jener sich als $(x -)y$ darstellt. Wesentlich ist hier, daß nicht etwa für x und y gleichzeitig Werte eingesetzt zu denken sind, sondern zunächst allein für x etwa der Wert a, wodurch als Zwischenstufe erst die Funktion mit dem Wert $a - y$ (kurz: die Funktion $a -$) entsteht, derart, daß nun erst die Ersetzung des y etwa durch den festen Wert b angängig wird.

fx ist also im obigen Falle der Wert einer Funktion, die nach Einsetzung eines Wertes für x noch nicht ein Ding des Grundbereiches (falls ein solches als Wert von $F(x, y)$ gemeint war), sondern wiederum eine Funktion liefert, deren Argument nunmehr y ist; d. h. f ist eine Funktion, deren Argument keiner Einschränkung zu unterliegen braucht, deren Funktionswert aber wieder eine Funktion ist. Die oben beschriebene Umformung werden wir von nun an für Funktionen von mehr als einer Veränderlichen stets durchführen bzw. durchgeführt denken, so daß diese durchweg in der Form

$$fxyz\ldots$$

erscheinen, die, wie schon gesagt, als Abkürzung für

$$(((fx)y)z)\ldots$$

gedeutet werden soll.

§ 3.

Es soll nunmehr eine Reihe von *individuellen Funktionen* von sehr allgemeiner Natur eingeführt werden. Ich nenne sie: die Identitätsfunktion I, die Konstanzfunktion C, die Vertauschungsfunktion T, die Zusammensetzungsfunktion Z und die Verschmelzungsfunktion S.

1. Unter der *Identitätsfunktion I* soll diejenige völlig bestimmte Funktion verstanden werden, deren Argumentwert keiner Einschränkung unterworfen ist und deren Funktionswert stets mit dem Argumentwert übereinstimmt, durch die also jedes Ding und jede Funktion sich selbst zugeordnet wird. Sie ist somit definiert durch die Gleichung

$$Ix = x,$$

in welcher das Gleichheitszeichen nicht etwa als logische Äquivalenz im Sinne der im logischen Aussagenkalkül üblichen Definition zu lesen ist, sondern besagt, daß die Ausdrücke links und rechts dasselbe bedeuten, d. h. daß der Funktionswert Ix stets derselbe ist wie der Argumentwert x, was man auch für x einsetzen mag. (So wäre z. B. $II = I$.)

2. Nunmehr sei der Argumentwert wieder ohne Einschränkung beliebig, während der Funktionswert unabhängig von jenem stets der feste Wert a sein soll. Diese Funktion ist ihrerseits von a abhängig, also von der Form Ca. Daß ihr Funktionswert stets a ist, wird geschrieben:

$$(Ca)y = a.$$

Und, indem wir nun auch a variabel lassen, erhalten wir:

$$(Cx)y = x \quad \text{bzw.} \quad Cxy = x$$

als Definitionsgleichung der *Konstanzfunktion C*. Diese Funktion C ist augenscheinlich von der auf S. 308 betrachteten Art; sie liefert nämlich erst durch Einsetzen eines festen Wertes für x eine Funktion mit dem Argument y. In der praktischen Anwendung leistet sie uns den Dienst, daß sie eine Größe x als „blinde" Veränderliche einzuführen gestattet.

3. Einen Ausdruck

$$fxy$$

kann man offenbar auch umgekehrt stets als aus

$$F(x, y)$$

entstanden betrachten, wo F durch das gegebene f eindeutig bestimmt ist. Schreibt man diesen Ausdruck nun andererseits um in

$$g\,y\,x,$$

indem man also y als Parameter betrachtet, so ist auch diese neue Funktion durch F und daher mittelbar auch durch f eindeutig gegeben.

Wir können daher die Funktion g als den Wert einer Funktion T für den Argumentwert f auffassen. Diese *Vertauschungsfunktion* T hat als Argument eine Funktion der Form $\varphi\,x\,y$, und der Funktionswert

$$\psi = T\varphi$$

ist diejenige Funktion $\psi\,x\,y$, für welche der Wert $\psi\,x\,y$ mit $\varphi\,y\,x$ bei allen den Argumentwerten x, y übereinstimmt, für die $\varphi\,y\,x$ einen Sinn hat. Wir schreiben diese Definition kurz:

$$(T\varphi)\,x\,y = \varphi\,y\,x,$$

wo die Klammern wiederum auch fehlen dürfen.

Die Funktion T bietet die Möglichkeit, die Reihenfolge der Glieder eines Ausdrucks abzuändern, und hilft damit über den Mangel des kommutativen Gesetzes bis zu einem gewissen Grade hinweg.

4. Erscheint an der Argumentstelle einer Funktion f der (von x abhängige) Wert einer anderen Funktion g, so hängt

$$f(g\,x)$$

augenscheinlich ebenfalls von x ab und kann folglich als der Wert einer dritten, durch f und g eindeutig bestimmten Funktion F betrachtet werden. In der Analysis spricht man hier bekanntlich ungenau von einer „Funktion von einer Funktion" — es müßte richtig heißen: eine Funktion von einem Funktionswert — und bezeichnet F als die aus f und g „zusammengesetzte" Funktion. Die Funktion F ist somit ihrerseits der Wert einer bestimmten Funktion Z' von f und g.

Wir könnten daher definieren:

$$[Z'(\varphi, \chi)]\,x = \varphi(\chi\,x).$$

Doch werden wir es im Sinne unserer früheren Verabredung vorziehen, Z' durch die zugehörige Funktion eines Arguments zu ersetzen, und erhalten demzufolge als Definitionsgleichung der *Zusammensetzungsfunktion* Z:

$$Z\varphi\,\chi\,x = \varphi(\chi\,x).$$

Vermittels der Funktion Z können Klammern innerhalb eines umfassenderen Ausdrucks verschoben werden (nicht eigentlich beseitigt, da sie stets noch hinzuzudenken sind); sie wirkt also im Sinne des hier ebenfalls nicht erfüllten assoziativen Gesetzes.

5. Setzt man in

$$f x y$$

für y den Wert einer Funktion g ein, und zwar genommen für dasselbe x, das als Argument von f auftritt, so kommt man auf einen Ausdruck

$$f x (g x)$$

oder, wie wir für den Augenblick etwas übersichtlicher schreiben wollen:

$$(f x)(g x).$$

Dies ist natürlich der Wert einer Funktion von x allein, also

$$(f x)(g x) = F x,$$

wo

$$F = S'(f, g)$$

wieder in einer völlig bestimmten Weise von den gegebenen Funktionen f und g abhängt. Wir haben demgemäß:

$$[S'(\varphi, \chi)] x = (\varphi x)(\chi x)$$

oder, nach der auch im vorigen Fall verwendeten Umformung:

$$S \varphi \chi x = (\varphi x)(\chi x)$$

als Definitionsgleichung der *Verschmelzungsfunktion* S.

Es wird gut sein, diese Funktion durch ein praktisches Beispiel dem Verständnis näherzubringen. Nehmen wir etwa für $f x y$ den Wert ${}^{x}\log y$ (d. h. den Logarithmus von y zu der Basis x) und für $g z$ den Funktionswert $1 + z$, so ergibt sich $(f x)(g x)$ augenscheinlich als ${}^{x}\log(1 + x)$, d. h. als der Wert einer Funktion von x, die mit den beiden gegebenen Funktionen eben durch unsere allgemeine Funktion S eindeutig verknüpft ist.

Der praktische Nutzen der Funktion S besteht ersichtlich darin, daß sie es ermöglicht, mehrmals auftretende Veränderliche — und bis zu einem gewissen Grade auch individuelle Funktionen — nur einmal auftreten zu lassen.

§ 4.

Es wird sich für die Durchführung unseres logisch-symbolischen Problems als belangreich erweisen, daß die oben erklärten fünf individuellen Funktionen I, C, T, Z, S des Funktionenkalküls nicht voneinander unabhängig sind, vielmehr zwei von ihnen, nämlich C und S, hinreichen, um die übrigen durch sie zu definieren. Und zwar bestehen hier die folgenden Zusammenhänge:

1. Es ist gemäß der Erklärung der Funktionen I und C:

$$I x = x = C x y.$$

Da y willkürlich ist, können wir dafür ein beliebiges Ding oder eine beliebige Funktion einsetzen, also z. B. Cx. Dies gibt:

$$Ix = (Cx)(Cx).$$

Nach der Erklärung von S bedeutet dies aber:

$$SCCx,$$

so daß wir erhalten:

$$I = SCC.\,^{3})$$

Übrigens kommt es in dem Ausdruck SCC auf das letzte Zeichen C gar nicht einmal an. Setzen wir nämlich oben für y nicht Cx, sondern die willkürliche Funktion φx, so ergibt sich entsprechend:

$$I = SC\varphi,$$

wo also für φ jede beliebige Funktion eingesetzt werden kann[4]).

2. Nach der Erklärung von Z ist

$$Zfgx = f(gx).$$

Weiter ist vermöge der bereits verwendeten Umformungen

$$f(gx) = (Cfx)(gx) = S(Cf)gx = (CSf)(Cf)gx.$$

Verschmelzung nach f ergibt:

$$S(CS)Cfgx,$$

also

$$Z = S(CS)C.$$

3. Ganz entsprechend läßt sich

$$Tfyx = fxy$$

weiter umformen in:

$$fx(Cyx) = (fx)(Cyx) = Sf(Cy)x = (Sf)(Cy)x = Z(Sf)Cyx$$
$$= ZZSfCyx = (ZZSf)Cyx = (ZZSf)(CCf)yx = S(ZZS)(CC)fyx.$$

Es gilt somit:

$$T = S(ZZS)(CC).$$

Setzt man hier für Z den oben gefundenen Ausdruck ein, so ist damit T ebenfalls auf C und S zurückgeführt.

§ 5.

Wir wollen nunmehr unsere Ergebnisse auf den besonderen Fall des Logikkalküls anwenden, in welchem die Grundelemente die Individuen und

---

[3]) Diese Zurückführung wurde mir von Herrn Boskowitz mitgeteilt, die etwas weniger einfache $(SC)(CC)$ bereits früher von Herrn Bernays.

[4]) Freilich nur eine solche, die für jedes x einen Sinn hat.

die Funktionen die Aussagefunktionen sind. Wir brauchen zunächst eine weitere individuelle Funktion, die diesem Kalkül eigentümlich ist. Der Ausdruck

$$fx\,|^{x}\,gx,$$

wo f und g Aussagefunktionen eines Arguments sind — auf solche dürfen wir uns gemäß einer früheren Bemerkung beschränken —, ist augenscheinlich eine bestimmte Funktion der beiden Funktionen f und g, also von der Form $U(f, g)$ oder, nach unserem Umformungsprinzip, Ufg. Damit haben wir

$$Ufg = fx\,|^{x}\,gx,$$

wo f und g nun natürlich Aussagefunktionen sind, als Definitionsgleichung der *Unverträglichkeitsfunktion* U.

Es besteht nun die bemerkenswerte Tatsache, daß jede logische Formel sich allein durch unsere individuellen Funktionen I, C, T, Z, S, U, also insbesondere schon durch C, S und U ausdrücken läßt.

Zunächst einmal läßt sich jede logische Formel vermittels der verallgemeinerten Strichsymbolik ausdrücken, wobei die gebundenen Veränderlichen (apparent variables) an den oberen Enden der Striche stehen. Dies gilt ohne Einschränkung, also für beliebige Aussagenordnungen und auch, wenn Beziehungen auftreten. Weiterhin läßt sich schrittweise mit geeigneter Verwendung der übrigen konstanten Funktionen an Stelle des Strichsymbols die Funktion U einführen.

Der Nachweis soll hier nicht vollständig durchgeführt, sondern nur die Rolle der verschiedenen individuellen Funktionen bei dieser Zurückführung erläutert werden.

Vermöge der Funktion C kann man erreichen, daß die beiden links und rechts vom Strich stehenden Ausdrücke Funktionen desselben Arguments sind.

So wäre z. B. der von f, g und y abhängige Ausdruck

$$fx\,|^{x}\,gy,$$

wo also rechts x nicht vorkommt, als

$$fx\,|^{x}\,C(gy)x$$

umzuschreiben. Kommt dagegen x rechts an anderer Stelle vor, so läßt es sich vermittels der Funktion T an den Schluß bringen, wobei es gegebenenfalls vermöge der Funktion Z aus Klammern befreit und, falls es mehrmals vorkommen sollte, vermöge der Funktion S verschmolzen werden muß. So haben wir z. B.:

$$fx\,|^{x}\,gxy = fx\,|^{x}\,Tgyx = Uf(Tgy).$$

Oder, um ein etwas verwickelteres Beispiel zu nehmen:

$$(fxy\,|^y\,gxy)\,|^x\,(hxz\,|^z\,kxz) = U(fx)(gx)\,|^x\,U(hx)(kx).$$

Hier ist z. B. der Ausdruck vor dem Strich in der folgenden Weise weiter zu behandeln:

$$U(fx)(gx) = ZUfx(gx) = S(ZUf)gx.$$

Der Gesamtausdruck wird damit:

$$S(ZUf)gx\,|^x\,S(ZUh)kx = U[S(ZUf)g][S(ZUh)k].$$

Wären im letzten Beispiel f und g identisch, so würden wir auf einen Ausdruck

$$S(ZUf)f$$

kommen. Um hier die Verschmelzung nach f durchführen zu können, bedienen wir uns der Funktion I, indem wir weiter rechnen:

$$S(ZUf)f = S(ZUf)(If) = [ZS(ZU)f](If) = S[ZS(ZU)]If.$$

Als praktisches Beispiel für die Behauptung dieses Paragraphen behandeln wir die folgende Aussage: „Zu jedem Prädikat gibt es ein mit ihm unverträgliches", d. h. „Zu jedem Prädikat f gibt es ein Prädikat g, so daß die Aussage $fx \,\&\, gx$ für kein Ding x richtig ist".

In der Hilbertschen Symbolik schreibt sich der Satz:

$$(f)(Eg)(x)\overline{fx \,\&\, gx}.$$

Dies wird zunächst:

$$(f)(Eg)(fx\,|^x\,gx)$$

und, indem man das partikuläre Urteil als Verneinung eines allgemeinen schreibt:

$$(f)\overline{(g)\overline{fx\,|^x\,gx}} \quad \text{bzw.} \quad (f)\overline{(g)\overline{fx\,|^x\,gx \,\&\, fx\,|^x\,gx}}.$$

Dies ist:

$$(f)\overline{(fx\,|^x\,gx)\,|^g\,(fx\,|^x\,gx)}.$$

Verfährt man entsprechend auch für f, so ergibt sich weiterhin:

$$\overline{(f)\,(fx\,|^x\,gx)\,|^g\,(fx\,|^x\,gx) \,\&\, (fx\,|^x\,gx)\,|^g\,(fx\,|^x\,gx)}$$
$$= [(fx\,|^x\,gx)\,|^g\,(fx\,|^x\,gx)]\,|^f\,[(fx\,|^x\,gx)\,|^g\,(fx\,|^x\,gx)].$$

Nunmehr erscheint das Strichsymbol als einziges logisches Verknüpfungszeichen. Führen wir jetzt die Unverträglichkeitsfunktion U ein, so erhalten wir zunächst:

$$[(Ufg)\,|^g\,(Ufg)]\,|^f\,[(Ufg)\,|^g\,(Ufg)]$$

und weiterhin:

$$[U(Uf)(Uf)]\,|^f\,[U(Uf)(Uf)].$$

Nun ist aber:

$$U(Uf)(Uf) = (ZUUf)(Uf) = S(ZUU)Uf,$$

womit der obige Ausdruck übergeht in:

$$[S(ZUU)Uf]|^f[S(ZUU)Uf],$$

d. h. aber:

$$U[S(ZUU)U]\,[S(ZUU)U].$$

§ 6.

Weiter als bis zu den Symbolen C, S und U läßt sich, soviel wir sehen, die Zurückführung nicht ohne Zwang treiben.

Rein schematisch könnte man freilich C, S und U sogar durch eine einzige Funktion ersetzen, indem man die neue Funktion J einführte durch die Festsetzung:

$$JC = U, \quad JS = C, \quad Jx = S,$$

wo x jedes von C und S verschiedene Ding ist. Wir stellen zunächst fest, daß J seinerseits von C und S verschieden ist, da nämlich J nur drei, C ebenso wie S dagegen unendlich viele Funktionswerte annimmt. Wir haben infolgedessen:

$$JJ = S, \quad J(JJ) = JS = C, \quad J[J(JJ)] = JC = U,$$

womit die Zurückführung in der Tat geleistet ist. Doch hat diese wegen ihrer augenscheinlichen Willkür wohl kaum sachliche Bedeutung.

Dagegen[5]) kann man sich auf einem anderen, natürlicheren Wege wenigstens von dem Zeichen U in gewissem Sinne befreien. Jede logische Formel enthält gewiß das Zeichen U und läßt sich, ganz wie wir dies früher für ein beliebiges Symbol überlegt haben, vermittels der individuellen Funktionen des allgemeinen Funktionenkalküls, also insbesondere vermittels C und S so umformen, daß U als Argument des gesamten Ausdrucks erscheint, dieser also die Form FU annimmt, wo F das U seinerseits nicht mehr enthält. Läßt man beim Hinschreiben das U als selbstverständlich weg, so kommt man in der Tat mit C und S aus.

Andererseits könnte man, mit Verzicht auf die äußerste Zurückführung der Grundfunktionszeichen, die Forderung aufstellen, daß die Klammern ganz vermieden werden sollen. Gehen wir von der Form FU aus, so läßt sich F allein vermöge Z so schreiben, daß alle Klammern verschwinden. Vermittels C, Z und S läßt sich also jede logische Formel als eine einfache Aufeinanderfolge dieser Zeichen ohne Klammern schreiben, mithin durch eine Zahl des triadischen Systems erschöpfend kennzeichnen.

---

[5]) Die folgenden Überlegungen rühren vom Bearbeiter her.

Was die Frage nach der *Eindeutigkeit* der betrachteten Zurückführung betrifft, so kann, rein symbolisch betrachtet, von einer solchen natürlich keine Rede sein, da jede Formel des alten wie auch des neuen Kalküls sich in mannigfacher Weise umformen läßt. Gleichwohl kann man in einem gewissen eingeschränkteren Sinne hier doch eine Eindeutigkeit des Entsprechens feststellen. Nennt man nämlich „gleichwertig" einerseits solche Formeln des alten Kalküls, die allein auf Grund von Definitionen, d. h. also ohne Benutzung der logischen Axiome — in denen natürlich die verallgemeinerte Sheffersche Verknüpfung nunmehr als Grundverknüpfung zu gelten hätte — aufeinander zurückgeführt werden können, und andererseits solche, die sich allein in den Typen der auftretenden Veränderlichen unterscheiden, so entsprechen in der Tat ein und derselben Formel des neuen Kalküls und ebenso einer jeden, die sich durch symbolische Rechnung aus ihr gewinnen läßt, alle und nur solche Formeln des alten Kalküls, die in dem eben erklärten Sinne untereinander gleichwertig sind. Die hier betrachtete Zurückführung der logischen Formeln hat also die bemerkenswerte Eigentümlichkeit, von den Axiomen der Logik unabhängig zu sein.

(Eingegangen am 15. 3. 1924.)